

Document downloaded from:

[\[http://redivia.gva.es/handle/20.500.11939/6686\]](http://redivia.gva.es/handle/20.500.11939/6686)

This paper must be cited as:

[Butelli, E., Garcia-Lor, A., Licciardello, C., Las Casas, G., Hill, L., Recupero, G. R., ... & Deng, X. (2017). Changes in anthocyanin production during domestication of Citrus. *Plant Physiology*, 173(4), 2225-2242.]

ivia
Institut Valencià
d'Investigacions Agràries

The final publication is available at

[\[http://dx.doi.org/10.1104/pp.16.01701\]](http://dx.doi.org/10.1104/pp.16.01701)

Copyright [ASPB]

Changes in anthocyanin production during domestication of *Citrus*

Eugenio Butelli¹, Andrés Garcia-Lor², Concetta Licciardello³, Giuseppina Las Casas⁴, Lionel Hill¹, Giuseppe Reforgiato Recupero³, Manjunath L Keremane⁵, Chandrika Ramadugu⁶, Robert Krueger⁵, Qiang Xu⁷, Xiuxin Deng⁷, Anne-Laure Fanciullino⁸, Yann Froelicher⁹, Luis Navarro² and Cathie Martin¹

¹John Innes Centre, Norwich NR4 7UH, United Kingdom; ²Instituto Valenciano de Investigaciones Agrarias, Carretera Moncada a Náquera Km. 4,5, 46113-Moncada, Valencia, Spain; ³Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria - Centro di Ricerca per l'Agricoltura e le Colture Mediterranee, 95024 Acireale, Italy; ⁴Dipartimento di Agricoltura, Alimentazione e Ambiente, University of Catania, 95123 Catania, Italy; ⁵USDA-ARS National Clonal Germplasm Repository for Citrus and Dates, 1060 Martin Luther King Blvd, Riverside, CA 92507-5437, USA; ⁶University of California Riverside, Riverside, CA 92511, USA; ⁷Key Laboratory of Horticultural Plant Biology of Ministry of Education, Huazhong Agricultural University, Wuhan 430070, P.R. China; ⁸INRA, UR1115 PSH, Avignon, France; ⁹CIRAD, UMR AGAP, Station INRA, F-20230 San Giuliano, France.

Author contributions

E.B. and C.M. planned and designed the research.
E.B., C.M., A.G-L., C.L., G.L.C., L.H. performed experiments.
M.K., C.R., R.K., G. R-R., A-L.F., Y.F., L.N., Q.X. and X.D. provided plant material, DNA, information and images.
E.B. and C.M. wrote the manuscript with input and comments from all the other authors.

Funding Information

C.M., E.B., were supported by the Institute Strategic Program Understanding and Exploiting Plant and Microbial Secondary Metabolism (BB/J004596/1) from the BBSRC and C.M., E.B., C.L., G.L.C and G.R.R. were supported by the European Union FP7 ATHENA collaborative project (Grant Agreement 245121). AGL and LN were partially supported by grant AGL2011-26490, from the Ministry of "Economía y Competitividad" and RTA2015-00069-00-00, from the Ministry of Economía, Industria y Competitividad-INIA- "Fondo Europeo de Desarrollo Regional (FEDER)", Spain. This work also benefited from the networking activities within the European-funded COST ACTION FA1106 QualityFruit.

Author for correspondence:

Eugenio Butelli: eugenio.butelli@jic.ac.uk

Tel: +44 (0)1603 450000

Keywords: anthocyanins, citrus, domestication, hybridization, MYB transcription factors, phylogenetic analysis.

One-sentence summary

Ruby, a regulatory gene encoding a MYB transcription factor, is essential for anthocyanin production and differences in its activity determine most of the natural variation in pigmentation in *Citrus* and related genera.

Abstract

Mandarin (*C. reticulata*), citron (*C. medica*) and pummelo (*C. maxima*) are important species of the genus *Citrus* and parents of the interspecific hybrids that constitute the most familiar commercial varieties of citrus: sweet orange, sour orange, clementine, lemon, lime and grapefruit. Citron produces anthocyanins in its young leaves and flowers, as do species in genera closely related to *Citrus*, but mandarins do not and pummelo varieties that produce anthocyanins have not been reported. We investigated the activity of the *Ruby* gene, which encodes a MYB transcription factor controlling anthocyanin biosynthesis, in different accessions of a range of *Citrus* species and in domesticated cultivars. A *white* mutant of lemon lacks functional alleles of *Ruby*, demonstrating that *Ruby* plays an essential role in anthocyanin production in *Citrus*. Almost all the natural variation in pigmentation by anthocyanins in *Citrus* species can be explained by differences in activity of the *Ruby* gene, caused by point mutations, deletions and insertions of transposable elements. Comparison of the allelic constitution of *Ruby* in different species and cultivars also helps to clarify many of the taxonomic relationships in different species of *Citrus*, confirms the derivation of commercial varieties during domestication, elucidates the relationships within the subgenus *Papeda* and allows a new genetic classification of mandarins.

Introduction

Anthocyanins are phenolic compounds responsible for the red, blue and purple colours in most angiosperm plants. They are accumulated in different tissues, particularly in flower petals, fruit peel and leaves (Winkel-Shirley, 2001). The pathway of anthocyanin biosynthesis is well characterized and the structural genes that encode the core biosynthetic enzymes have been identified from many species (Tanaka et al., 2008). These genes are coordinately induced by a conserved regulatory complex formed by the interaction between MYB transcription factors, bHLH transcription factors and WD40-repeat proteins, called the MBW complex (Quattrocchio et al., 1993; Ramsay and Glover, 2005; Koes et al., 2005). DNA-binding specificity is provided principally by the MYB factors. Mutations in the MYB regulatory genes often result in lack of expression of the biosynthetic genes and loss of anthocyanin pigmentation, as exemplified by the occurrence of 'red' and 'white' varieties of grape (Kobayashi et al., 2004), apple (Espley et al., 2007) and species of petunia (Quattrocchio et al., 1999; Hoballah et al., 2007) and antirrhinum (Schwinn et al., 2006).

In *Citrus*, the ability to accumulate anthocyanins is not a universal feature (Fig. 1). In lemons and limes, purple pigmentation is clearly present in young leaves and flowers (new growth). In contrast, other cultivated *Citrus* varieties of economic importance are completely unable to synthesize anthocyanins (Hodgson, 1967). Sweet oranges, with the notable exception of the striking fruit-specific red-coloured, blood varieties, do not display any anthocyanin pigmentation.

The taxonomy of *Citrus*, a genus of plants belonging to the family *Rutaceae* (Supplemental Fig. S1), is complicated, as illustrated by the two most widely accepted classification systems, those proposed by Swingle and Tanaka, recognizing 16 and 162 species, respectively (Swingle, 1946; Swingle and Reece, 1967; Tanaka, 1977). In *Citrus*, the definition of a species is obscured by the ability of *Citrus* species to hybridize easily within the genus and also with closely related genera. Sexual promiscuity and thousands of years of human cultivation have generated many hybrids whose origin cannot be inferred solely on the basis of morphological and geographical data. However, a growing body of evidence supports the view that there are three basic or primary species: mandarin (*C.*

109 *reticulata*), pummelo (*C. maxima*) and citron (*C. medica*) (Scora, 1975; Barrett and
110 Rhodes, 1976). All the other 'species' are hybrids derived from these three primary
111 species (Moore, 2001). A fourth species belonging to the subgenus *Papeda* was
112 involved in the derivation of some limes (Nicolosi et al., 2000). The term 'secondary
113 species' has been used widely to define *Citrus* hybrids since most of them are
114 apomictic, polyembryonic and are normally propagated vegetatively by grafting. This
115 ensures that plants maintain their hybrid genetic constitution over generations but
116 also makes genetic studies challenging and crop improvement by conventional
117 breeding, almost impossible. Within a given species or hybrid, phenotypic
118 differences between cultivars are usually the result of spontaneous or induced
119 somatic mutations that have been selected and propagated.

120 We have previously shown that, in different blood varieties of sweet orange, the
121 insertion of retrotransposons in the promoter of *Ruby*, a regulatory gene encoding a
122 MYB transcription factor, is responsible for fruit-specific accumulation of
123 anthocyanins (Butelli et al., 2012). Here, our objective was to determine whether
124 *Ruby* controls the accumulation of anthocyanins in young leaves and flowers of
125 pigmented *Citrus* species and hybrids and whether mutations in this gene can
126 account for the inability of some *Citrus* varieties to develop pigmentation. We used
127 sequence information to determine whether the phylogeny of *Ruby* is congruent with
128 the generally accepted phylogeny of *Citrus* and the proposed ancestry of
129 commercially important hybrids.

132 Results

134 The *Ruby* locus in *Citrus* and related genera

135 Most taxonomists divide the group recognized as 'true *Citrus* fruit trees' into six
136 related genera: *Citrus*, *Poncirus*, *Microcitrus*, *Eremocitrus*, *Fortunella* and *Clymenia*
137 (Swingle and Reece, 1967; Supplemental Fig. S1). The genus *Citrus* is further
138 divided into two subgenera: *Citrus*, which includes the common cultivated plants with
139 commercially important fruit, and *Papeda* (Swingle, 1946; Krueger and Navarro,
140 2007).

141 We attempted a comprehensive analysis of the sequence variation at the *Ruby* locus

in the three primary *Citrus* species, in four members of the subgenus *Papeda*, in several interspecific hybrids and in accessions belonging to four genera related to *Citrus* that are relevant to understanding the regulation of anthocyanin production. For each, we sequenced the *Ruby* gene and at least 1.6 kb of the region upstream of the coding sequence. The species and hybrids we have studied are listed in Table 1. *Ruby* is involved in fruit-specific accumulation of anthocyanins in blood varieties of sweet orange (Butelli et al., 2012). A second gene, which we named *Ruby2*, also encodes a putative R2R3 MYB transcriptional activator with the potential to regulate anthocyanin biosynthesis. The *Ruby2* and *Ruby* genes are arranged in tandem on chromosome 6 and are separated by an intergenic region ranging from 1.6 to 12 kb in different species. In *Citrus*, all the *Ruby2* alleles are predicted to encode nonfunctional proteins because of the presence of different combinations of deletions, frameshift and stop mutations (Supplemental Data Set S1). No other genes with similarity to *Ruby* and encoding proteins belonging to subgroup 6 of the R2R3MYB family (known to regulate anthocyanin biosynthesis; Stracke et al., 2001) were detected in the high-quality reference haploid clementine genome, nor in two different genome annotation projects of sweet orange (<https://phytozome.jgi.doe.gov>; <http://citrus.hzau.edu.cn/orange/>). A previous genome-wide targeted analysis of *R2R3 MYB* genes in sweet orange (Liu et al., 2014) and examination of the genome sequence of other *Citrus* species (QX and XD personal communication) also failed to identify other paralogs of known anthocyanin-related *MYB* genes. Based on these observations, the hypothesis was developed that, in all the accessions belonging to the genus *Citrus*, *Ruby* is the only R2R3 MYB transcriptional activator involved in anthocyanin biosynthesis. Our objective was to determine whether variation within the *Ruby* locus is responsible for the differential ability of different accessions to produce anthocyanins.

***Ruby* in citron**

Citron (*C. medica*) is a true species with a very low level of heterozygosity. Most varieties are able to accumulate anthocyanins in young leaves and flowers (Fig. 1A), where cyanidin 3-glucoside and peonidin 3-(6"-malonyl)-glucoside accounted for approximately 80% of the total anthocyanin content (Fabroni et al., 2016). The three varieties of citron we tested (Table 1) contained the same *Ruby* allele,

predicted to be active and named R^C , in homozygous form. R^C was inherited, identical in sequence, in many hybrids including lemon where its expression was characterized.

Lying at position -902 relative to the start of transcription in R^C is a 5435 bp retrotransposon, which we named Tc13 (Supplemental Fig. S2A and S2B). Tc13 is 92 and 91% identical in sequence to the two retroelements that are responsible for anthocyanin accumulation in Mediterranean (Tcs1) and Chinese blood oranges (Tcs2) respectively (Butelli et al., 2012). All three elements are members of the same family of Copia-like LTR retrotransposons but, while Tcs1 and Tcs2 are intact, active and recently-inserted elements, Tc13 in citron is defective and is the result of an ancient retrotransposition event, as deduced by the divergence of LTR sequences that can be used to estimate the time of integration of LTR retrotransposons (SanMiguel et al., 1998). The two LTRs of Tc13 differ by seven nucleotides and a 10 bp InDel, and the open reading frame of Tc13 contains four stop mutations and one 5 bp deletion causing a frameshift mutation that would inactivate the encoded polyprotein (Supplemental Fig. S2B and S2C). These observations, together with its insertion at a considerable distance from the start of transcription, suggest that the presence of the Tc13 retrotransposon likely has little or no impact on *Ruby* expression.

***Ruby* in pummelo**

Pummelo (*C. maxima*) is a true species of *Citrus* with many cultivars and high genetic diversity. Anthocyanin pigmentation has never been reported in pummelo. The four accessions we initially considered ('Chandler', 'Reinking', 'Siamese acidless' and 'Kao Pan') contain different combinations of three *Ruby* alleles that we named r^{STOPp} , r^{TATA} and r^{CAAT} (Table 1).

The r^{STOPp} allele is characterized by the presence of a stop mutation (TGG to TGA at positions equivalent to +1481 in the R^C reference allele) in the third exon that would result in the truncation of the Ruby protein involving loss of its carboxy-terminal 116 amino acids. This mutation results in a complete loss of function of the Ruby protein as demonstrated by the lack of anthocyanin production when a truncated cDNA was transiently overexpressed in *Nicotiana benthamiana* (Fig. 2B). In contrast, the full-length *Ruby* cDNA induced the accumulation of visible levels of delphinidin-3-rutinoside in *N. benthamiana* (Fig. 2A).

Although the coding sequences of the other two alleles identified in pummelo encoded intact proteins, we detected a number of changes in the core promoter that might disrupt *Ruby* expression. The r^{TATA} allele contains a T>C change in the centre of the putative TATA box. This mutation lies at -32 bp relative to the start of transcription (as determined in lemon; see below). In the allele r^{CAAT} , we identified another T>C transition this time within a putative CAAT box (at position -73 bp relative to the start of transcription and also present in r^{STOPp}). TATA and CAAT boxes are *cis*-acting elements that determine the efficiency of the promoter in many eukaryotic genes (Cooper and Hausman, 2016) and the two mutations, specific for the pummelo accessions, lie in a region that is well conserved in *Citrus* and related genera (Supplemental Fig. S3A). Sweet orange carries the r^{TATA} allele from its pummelo parent (see below and Butelli et al., 2012) and retrotransposon insertions into this allele provide either a new TATA box or activate a cryptic TATA box to activate *Ruby* expression in blood oranges (Butelli et al., 2012).

To confirm the prediction that both r^{CAAT} and r^{TATA} alleles are not expressed, we analyzed ‘Chandler’ pummelo since both alleles are present in this variety (Table 1 and Supplemental Fig. S3B). We were unable to isolate *Ruby* cDNA fragments and we could detect no *Ruby* expression in leaves of ‘Chandler’ pummelo (Supplemental Fig. S12C).

Although, theoretically, the lack of expression of *Ruby* from the r^{TATA} and r^{CAAT} alleles could have resulted from mutations in a closely linked regulator of *Ruby*, these alleles failed to complement alleles of *Ruby* with interruptions in the *Ruby* coding sequence (such as r^{STOPp}) in pummelo and r^{TATA} failed to complement the alleles from mandarin (which carry large deletions in the *Ruby* gene) when combined in sweet orange (Butelli et al., 2012).

All four pummelo varieties analyzed contained at least one copy of a *Ruby* gene with an intact coding sequence. Therefore, pummelo appears to maintain the potential to produce anthocyanins, a potential that could be fulfilled by recombination of sequences lying between the mutated CAAT box allele and the mutated TATA box allele. Such an event is possible because pummelo, unlike the majority of cultivated *Citrus* plants, produces true zygotic, monoembryonic seeds. Although purple pummelo varieties have not been reported, an individual wild tree bearing purple/red-skinned fruit was identified by Prof. Xiuxin Deng in a mountainous region of Hubei province, China (Supplemental Fig. S3C). Sequence analysis indicated that this

pigmented variety carried a *Ruby* allele with both an 'active CAAT box' and an 'active TATA box' that might therefore be the result of such a recombination event. This unique allele, named R^P , was predicted to be fully active since an intact coding sequence was under the control of a core promoter without mutations in the CAAT or TATA boxes (Fig. 3 and Supplemental Fig. S3). Although there were mismatches in the sequence of the region upstream of the TATA box in the R^P alleles, this sequence bore a greater similarity to the equivalent regions of r^{TATA} than to r^{CAAT} , including the absence of a 22 bp deletion. Downstream of the active TATA box motif in R^P , the sequence differed from r^{TATA} and most of the SNPs present in the area sequenced were identical between R^P and r^{CAAT} . (Supplemental Fig. S4). This new active R^P allele supports the possibility that it might be the result of recombination. Although the occurrence of such a recombination event is likely to be rare, this particular individual might have been identified and preserved by the local population because of its coloured fruit peel. The phenotype of the R^P allele might reflect the expression pattern of *Ruby* in a pigmented ancestor of pummelo, before loss-of-function *Ruby* alleles began to predominate as a result of evolution or human selection.

***Ruby* in mandarin**

The mandarin group is composed of numerous varieties with considerable genetic diversity but treated as a single species under the name *C. reticulata* according to Swingle (Garcia-Lor et al., 2015).

We analyzed two common types of mandarin: 'Ponkan', an old variety thought to be a pure, prototypical *C. reticulata* (Hodgson, 1967) and 'Avena', a Mediterranean variety also known as 'Willowleaf'. The structure and the sequence of the *Ruby* locus was identical in the two varieties and carried one allele (r^{STOPm}) with the same stop mutation identified in pummelo and another (r^{DEL2}) with a 2 kb deletion involving the first two exons and a large portion of the upstream regulatory region of *Ruby* (Table 1, Fig. 3 and Supplemental Fig. S5). Sequence alignments suggested that the r^{STOPm} allele was derived from pummelo because the two alleles, r^{STOPm} and r^{STOPp} , showed 99% nucleotide identity and contained identical polymorphisms, including the distinctive stop mutation, that were not present in *Ruby* alleles from other *Citrus* or related species (Supplemental Fig. S5). Comparative sequencing of different *Citrus* varieties (Wu et al., 2014) has indicated a substantial admixture from pummelo in

‘Ponkan’ and ‘Willowleaf’ mandarins. In particular, the reported nucleotide heterozygosity in the terminal region of chromosome 6, where *Ruby* is located, supports the likely pummelo origin of a *Ruby* allele in some varieties of mandarin. Since an ancestral mandarin may have had a distinct allelic composition at the *Ruby* locus, we looked among existing mandarins for different alleles of *Ruby*. Tanaka recognized 36 mandarin species separated into five taxonomic groups (Tanaka, 1977). We sequenced the *Ruby* alleles from 17 mandarin varieties, representative of accessions from all five groups (Table 1).

On the basis of their constitution of *Ruby* alleles, these mandarin accessions could be divided into three groups (Fig. 4). Group A ($r^{DEL0.8}$, $r^{DEL0.8}$) was characterized by being homozygous for a deletion of ~0.8 kb that spans the first two exons of the *Ruby* gene (between positions -175 and +629 relative to the start of transcription in the reference allele R^C). Mandarins of group B (r^{DEL2} , r^{DEL2}) were homozygous for a larger deletion of ~2 kb (upstream of position +729 in R^C). Mandarins of group C (r^{DEL2} , r^{STOPm}) were polymorphic for *Ruby* alleles, containing one allele introgressed from pummelo (r^{STOPm}) and one allele with the 2 kb deletion (r^{DEL2}) also present in group B. This polymorphic genetic constitution has probably been maintained because group C varieties, including ‘Ponkan’ and ‘Willowleaf’, are highly apomictic. None of the alleles in mandarin encode an active *Ruby* protein, explaining why mandarins have never been associated with anthocyanin pigmentation.

The short deletion allele ($r^{DEL0.8}$) is highly polymorphic (Supplemental Fig. S5). Group A includes varieties of Japanese origin that show admixture from other species (Kitajima, 2007; Curk et al., 2015; Garcia-Lor et al., 2015). Our sequencing of genes other than *Ruby* in satsumas (*C. unshiu*) suggested introgression from pummelo that is different from that documented in ‘Ponkan’ and ‘Willowleaf’ (Wu et al., 2014). *Citrus tachibana*, also in group A, contained a *Ruby2* allele derived from an unknown species as defined by distinct SNPs.

Mandarins in group B include small-fruited sour mandarins (*C. sunki* and *C. reshni*) of reputed Chinese origin. Unlike $r^{DEL0.8}$, the *Ruby* allele of this group with the larger deletion (r^{DEL2}) did not show polymorphisms within or between accessions. The long 2 kb deletion encompassed the short one (0.8 kb), from which it may have been derived relatively recently.

Our data suggested that group C mandarins were derived from group B following hybridization with pummelo, an event that did not restore anthocyanin biosynthesis.

This hybridization may have been a significant step in mandarin evolution that introduced new features including fruit quality and general plant fitness. Based on the allelic composition of *Ruby*, many mandarin varieties are likely to be hybrids between members of the three subgroups (Fig. 4). The different combinations of the three distinct alleles could help identify the origin and classify mandarin varieties that tend to cluster together in phylogenetic studies. The allelic constitution at the *Ruby* locus also indicated the involvement of other *Citrus* species in important varieties usually classified as mandarins (Fig. 4). In particular, our analysis supports the conclusion that *C. nobilis* originated as a hybrid between mandarin (group A) and sweet orange (Hodgson, 1967). Within *C. nobilis*, the common accessions 'King' and 'Kunembo' inherited different *Ruby* alleles from their sweet orange parent, confirming that they had independent origins (Penjor et al., 2013) (Table 1 and Fig. 4). Our analysis also confirmed a major contribution of papeda (*C. micrantha*) to the 'Nasranan' mandarin *C. amblycarpa* (Froelicher et al., 2011). *Citrus indica*, which was considered by both Tanaka and Swingle to be a wild mandarin species, shows no relationship to any of the other mandarins and most likely is a true species distantly related to citron (Supplemental Fig. S6).

Ruby* in *Papeda

Most *Citrus* plants are hybrids derived from different combinations of the three true species of *Citrus* (pummelo, mandarin and citron). Contributions from a fourth species belonging to the subgenus *Papeda* have been demonstrated in some limes (Nicolosi et al., 2000). *Papeda* is considered to be an old and primitive type of *Citrus* but its monophyletic status and its subdivision from *Citrus* are debated (Federici et al., 1998; Bayer et al., 2009; Carbonell-Caballero et al., 2015). We analyzed four different species of *Papeda* (*C. micrantha*, *C. ichangensis*, *C. macroptera* and *C. hystrix*; Table 1). All of them were able to accumulate anthocyanins in new growth but to differing extents. In *C. ichangensis* pigmentation is strong (Fig. 1C) whereas in *C. micrantha* pigmentation is often not visible (Fig. 1D) or limited to a trace of purple on the outside of young buds (Hodgson, 1967). The *Ruby* alleles in *C. ichangensis* (R^{Pap1} , R^{Pap1}) and *C. micrantha* (R^{Pap2} , R^{Pap2}) were unrelated while, surprisingly, *C. macroptera* and *C. hystrix* appeared to be identical hybrids of *C. micrantha* and *C. ichangensis* (Table 1). All the *Ruby* alleles in *Papeda* had intact open reading frames, consistent with their pigmented phenotypes. Both alleles from *C. micrantha*,

however, had an insertion of a defective Copia-like retrotransposon (named Tcm7), lacking both LTRs and inserted in antisense orientation 0.3 kb upstream of the predicted start of transcription (Fig. 3). This insertion was also associated with a large deletion of the promoter region that extends to include part of the upstream *Ruby2* gene. The insertion of the retrotransposon and the deletion of part of the promoter are likely to affect the expression of *Ruby* and may explain the relatively weak pigmentation in new growth of *C. micrantha* compared to *C. ichangensis* (Fig. 1D and 1C respectively).

***Ruby* in *Citrus* hybrids**

Many commercial *Citrus* varieties have a hybrid origin (Scora, 1975; Barrett and Rhodes, 1976; Nicolosi et al., 2000; Moore, 2001). Hybrids derived from citron display intense purple pigmentation in young, developing leaves and flowers where anthocyanins can provide protection for vulnerable tissues undergoing new growth. In contrast, hybrids involving mandarin and pummelo have inherited only nonfunctional *Ruby* alleles and do not produce anthocyanins (Fig. 5 and Supplemental Fig. S7).

Citron is one parent of lemon and of several hybrids often collectively defined as 'limes'. Citron is usually fertilized by self-pollination and always served as the male parent in the formation of cultivated hybrids (Nicolosi et al., 2000; Curk et al., 2016). To determine the role of *Ruby* in pigmented hybrids, we focused on lemon (*C. limon*) because of the availability of a unique *white* mutant, 'Zagara Bianca', which is completely unable to accumulate anthocyanins and most probably arose as bud sport of the cultivar 'Femminello' (Cultrone et al., 2010). The anthocyanin composition in young leaves of the wild type accession was determined by UHPLC-ESI-MS. Two compounds, cyanidin 3-(6"-malonyl)- β -glucoside and cyanidin 3-glucoside, accounted for most of the anthocyanin content (Fig. 6A and Supplemental Fig. S8, S9 and S10). The same compounds are the main anthocyanins identified in fruit of blood varieties of sweet orange and in new growth of many types of citrus cultivars, including lemon (Hillebrand et al., 2004; Fabroni et al., 2016). No anthocyanins could be detected in the *white* mutant. *Ruby*, was expressed in pigmented young flowers and leaves of the wild type accession but not in the corresponding tissues of the mutant (Fig. 6B).

The two lemon accessions were compared by Southern blot analysis. Genomic

DNA, hybridized to a probe of the full-length *Ruby* gene (Fig. 6C), displayed two distinct *Hind*III bands in wild type lemon, one of which was clearly absent in the mutant (Fig. 6D). The common band present in both accessions corresponded to the *Ruby* allele r^{STOPp} carrying the nonsense mutation that disrupts protein function (Fig. 6C). This nonfunctional allele was inherited from pummelo via sour orange (Fig. 5). The allele R^C , derived from citron, was missing in the *white* mutant, showing that *Ruby* is essential for anthocyanin biosynthesis in lemon and, by inference, in other species of *Citrus*.

The active allele of *Ruby*, R^C , encodes a transcript which translates into a predicted protein of 262 amino acids with only eight differences compared to the active *Ruby* protein in blood varieties of sweet orange. The start of *Ruby* transcription (designated +1) in lemon was mapped using 5' RACE. A TATA box and a potential CAAT box were identified at positions -34 and -74 respectively. The *Ruby* gene contains three exons at positions +103 to +220, +319 to +447 and +1243 to +1783. Citron was involved in the independent origin of different varieties of lime. The most common are Mexican lime (*C. aurantifolia*) and Persian lime (*C. latifolia*). The composition of *Ruby* alleles in Mexican lime (R^{Pap2} , R^C) was in agreement with papeda (*C. micrantha*) being its female parent (Nicolosi et al., 2000). Persian lime (r^{STOPp} , R^C) displayed the same *Ruby* alleles as lemon, which is compatible with it having been derived from a cross between lemon and Mexican lime as proposed by Bayer et al., (2009) based on chloroplast DNA marker analysis and Curk et al., (2016) using nuclear and cytoplasmic marker analyses (Fig. 5).

The origin of other citron-derived hybrids of minor commercial importance has received little attention. Palestine sweet lime (*C. limettioides*), carried the $r^{DEL0.8}$ allele in all the accessions analyzed (including the acidic form 'Soh Synteng'; Table 1) which argues against the contribution of either sweet or sour orange as suggested by Nicolosi et al. (2000) and rather suggests the presence of a pummelo x mandarin hybrid in its parentage as proposed by Curk et al. (2014 and 2016) based on haplotyping, nuclear and cytoplasmic DNA marker analyses. This hybrid, yet to be identified, might also have been involved in the derivation of Meyer lemon, which also contains the $r^{DEL0.8}$ allele. The composition of *Ruby* alleles (r^{DEL2} , R^C) in limetta (*C. limetta*) was consistent with the involvement of sour orange and citron as parents (Curk et al., 2014 and 2016). For Rangpur lime (*C. limonia*), the composition of *Ruby* alleles (r^{DEL2} , R^C) suggested a mandarin (group B) x citron origin.

The absence of anthocyanin pigmentation in flowers and leaves of many commercially important hybrids is associated with the absence of citron in their parentage (Fig. 5). For example, sour orange (*C. aurantium*), important for the production of marmalade and also used widely as a rootstock in the past, is a direct pummelo x mandarin F1 hybrid (Wu et al., 2014), confirmed by its composition of *Ruby* alleles (r^{STOPp} , r^{DEL2}) and never produces anthocyanins. Grapefruit (*C. paradisi*) was obtained in the West Indies in the 18th century, probably as a natural hybrid between pummelo and sweet orange (Morton, 1987; Moore, 2001). This origin is supported by its *Ruby* alleles (r^{STOPp} , r^{DEL2}) and grapefruit never produces anthocyanins. Clementine (*C. clementina*) is considered to have arisen as a natural hybrid of sweet orange and Mediterranean mandarin based on morphological markers and sequence data (Hodgson, 1967; Nicolosi et al., 2000; Wu et al., 2014). Its lack of anthocyanins is consistent with its composition of *Ruby* alleles (r^{STOPm} , r^{DEL2}) and this hybrid origin. Sweet orange (*C. sinensis*) is considered a complex hybrid involving mandarin (providing the predominant genetic material) and pummelo (Garcia-Lor et al., 2013b; Xu et al., 2013; Wu et al., 2014) and the allelic composition of the *Ruby* locus (r^{TATA} , r^{DEL2}) confirmed this. We have been unable to detect any expression of *Ruby* in different tissues of common varieties of sweet orange (Butelli et al., 2012), consistent with the complete absence of anthocyanin pigmentation. As discussed for pummelo, from which it is derived, the allele r^{TATA} in sweet orange is characterized by a T>C transition within the TATA box that prevents the gene from being expressed. In two independent blood varieties of sweet orange, however, the insertion of transposable elements has provided either a new TATA box and transcriptional start site (Tcs1), or promoted the use of a cryptic alternative start site (Tcs2) upstream of the T>C mutation (Supplemental Fig. S2A and S2D; Butelli et al., 2012). Tcs1 and Tcs2 also provide new instructions for cold-dependent, fruit-specific expression of *Ruby* but are unable to restore the normal pattern of anthocyanin accumulation in new growth of leaves and flowers. These data confirm that r^{TATA} fails to express *Ruby* because it has a defective TATA box.

Allelic constitution of *Ruby* in genera related to *Citrus*

In addition to *Citrus*, we considered three distinct genera - *Poncirus*, *Microcitrus* and *Fortunella* - that are classified within the 'true *Citrus* fruit trees' and a more distant

genus, *Severinia*, considered a 'primitive *Citrus* fruit tree' (Swingle and Reece, 1967; Supplemental Fig. S1). These genera are relevant in the context of anthocyanin pigmentation.

Trifoliate orange (*Poncirus trifoliata*) is a species, unusually deciduous, that is commonly used as a rootstock for *Citrus* scions. Young and senescent *P. trifoliata* leaves are pigmented with anthocyanins (Supplemental Fig. S11A). *Ruby* is predicted to be functional in *P. trifoliata* (Fig. 3) and the two alleles we identified (R^{PonA} and R^{PonB}) were very similar in sequence and both were expressed in pigmented leaves (Supplemental Fig. S11B and S11C).

In *Microcitrus*, we initially analyzed a variety (*M. australasica* var 'Sanguinea'), native to Australia, where it arose spontaneously (Netzel et al., 2007). This variety accumulates high levels of anthocyanins not only in new growth but also in fruit peel and pulp. We identified two *Ruby* alleles, R^{Mic1} and R^{Mic2} , that could be distinguished by the presence of a 0.5 kb deletion in the promoter (between positions equivalent to -887 to -364 in the R^C reference allele) of the R^{Mic1} allele or a subterminal in-frame 81 bp deletion (between positions +1650 and +1730) in the coding region of the R^{Mic2} allele (Fig. 3). In different tissues, the expression of *Ruby* was associated with anthocyanin pigmentation (Supplemental Fig. S12). Interestingly, while both alleles were expressed in leaves, in fruit we could detect only monoallelic expression of R^{Mic2} (Supplemental Fig. S12E), suggesting that the subterminal deletion, affecting the aminoacid sequence of a region that is not conserved in anthocyanin-specific *Myb* regulatory genes, is unlikely to affect protein function. To confirm the functionality of the transcription factor encoded by R^{Mic2} , we transiently overexpressed the corresponding cDNA in leaves of *N. benthamiana*. As shown in Fig. 2C, the construct was able to induce anthocyanin biosynthesis fully with no significant differences compared to the full-length *Ruby* cDNA isolated from 'Moro' blood orange.

Further analysis of the promoter region of R^{Mic2} revealed the presence of an insertion of a 0.4 kb repetitive element with an accompanying 7.5 kb deletion involving the entire *Ruby2* gene and a translocation of 0.5 kb of DNA from chromosome 5 at a position equivalent to -856 in the R^C reference allele (Fig. 3). The insertion, flanked by target duplication sites, corresponded to a repetitive element containing consensus motifs for SINE nonautonomous retrotransposons. Although it was difficult to assess the relative effects of the insertion, the deletion and the

translocation on the expression of the downstream *Ruby* gene, the presence of this rearrangement provided further evidence of the importance of transposable elements in reshaping gene structure and expression. This was confirmed by the analysis of a different variety of *M. australasica* with anthocyanin production in new growth but green-yellow fruit that do not accumulate anthocyanins. In this accession (Table 1), the complex mutation in the promoter of *Ruby* was not present and both alleles corresponded to R^{Mic1} .

We analyzed two ornamental species belonging to the genus *Fortunella*: ‘Nagami’ oval kumquat (*F. margarita*), and ‘Malayan’ kumquat (*F. polyandra*) which do not accumulate anthocyanins and found only nonfunctional *Ruby* alleles (r^{For}) in both species, all carrying a single nucleotide insertion at position +444 in the second exon that resulted in a frameshift mutation (Fig. 3). Calamondin (*Citrofortunella microcarpa*), another common variety that is anthocyaninless, also contained this defective allele from *Fortunella* (r^{For}) together with the $r^{DEL0.8}$ allele. This allelic composition confirmed the presence of mandarin in its ancestry but argued against a contribution by the sour variety ‘Sunki’, frequently suggested as a candidate parent (Lai and Chen, 2008).

We sequenced *Ruby* in Chinese box orange (*Severinia buxifolia*) because this species is a distant relative of the genus *Citrus* that is often used as an outgroup in phylogenetic studies. *Severinia buxifolia* produces highly pigmented, small fruit that become black at maturity. Both *Ruby* alleles were characterized by a repetitive SINE-like element of 0.5 kb inserted 0.4 kb upstream of the predicted start of transcription (Fig. 3). As in *M. australasica* var ‘Sanguinea’, it is possible that the presence of a SINE-like elements could influence the strong anthocyanin pigmentation in fruit.

Phylogenetic analysis

We conducted a phylogenetic analysis of 18 *Ruby* alleles representing the diversity of *Citrus* and related genera. Alleles from primary species only were used in the analysis because the *Ruby* sequences in the hybrids are identical to those identified in the primary parental species (Fig. 5).

Among all of the models tested via the Phylemon website (Sánchez et al., 2011), the one with the best fit was HKY+G+F. The phylogenetic relationships between *Citrus* species and their relatives inferred from the maximum-likelihood (ML) method using

this model are shown in Fig. 7.

The model that best fitted the data using the MEGA (Molecular Evolutionary Genetics Analysis) software was the T92 model (Kimura, 1980; Tamura, 1992; Tamura et al., 2011). The phylogenetic relationships between *Citrus* species and their relatives inferred using MEGA software are shown in Supplemental Fig. S13. Phylemon and MEGA phylogenies suggested the same relationships between *Ruby* alleles indicating a clear separation between *Citrus* and related genera (strong branch support). *Poncirus* and *Microcitrus* were quite distant from *Citrus* (high bootstrap values) while *Fortunella* was more closely related. This analysis supported the phylogenetic relationships in the *Citrinae* established using other markers (Nicolosi et al., 2000; Garcia-Lor et al., 2013a and 2013b; Ramadugu et al., 2013; Penjor et al., 2013) but did not reveal the distinct phylogenetic cluster of citron and *Microcitrus* established using the complete chloroplast DNA of 34 species (Carbonell-Caballero et al., 2015).

A division between the two subgenera, *Citrus* and *Papeda*, was not supported by our analysis and has been challenged in other studies (Penjor et al., 2010; Carbonell-Caballero et al., 2015). Instead our data suggested a polyphyletic origin of the *Papeda* group. The *Ruby* alleles in *C. micrantha* are related to those of *C. medica* and *C. indica* while the *Ruby* alleles of *C. ichangensis* clustered with the alleles from mandarin. We propose that the subgenus *Papeda* contains DNA from at least two distinct unrelated species (*C. micrantha* and *C. ichangensis*) and several hybrids between them (such as *C. hystrix* and *C. macroptera*). This distinction could explain much of the conflicting data on the origin and classification of *Papeda*. Whole genome sequencing would confirm or refute these suggestions.

The analysis in Fig. 7 also confirmed that r^{STOPm} , one of the alleles identified in common mandarin varieties such as 'Ponkan' and 'Willowleaf', lies within the clade of pummelo alleles and has been introgressed into mandarin from pummelo as suggested by Wu et al. (2014).

Discussion

Anthocyanins are widely distributed plant pigments with a diverse range of functions

and tissue distributions. Since they are not essential for plant survival, they provide a unique visual marker for the occurrence of mutations and activation of transposable elements, which may result in changes to the patterns of plant pigmentation as demonstrated in maize, antirrhinum, petunia and blood varieties of sweet orange (Malinowski, 1935; Fincham, 1967; McClintock, 1984; Coen et al., 1986; Butelli et al., 2012).

Ruby is a regulatory gene encoding a transcription factor that is essential for anthocyanin accumulation in *Citrus* and related genera. The *Ruby* locus contains two potential *MYB* regulators of anthocyanin biosynthesis (*Ruby* and the inactive *Ruby2*). Many angiosperm plants have experienced extensive duplication of these regulatory genes. The amplification of *MYB* genes belonging to subgroup 6 that regulate anthocyanin biosynthesis may be a general mechanism underlying the diversity of floral pigmentation patterns as observed in petunia (Bombarely et al., 2016) and antirrhinum (Schwinn et al., 2006) but might also be the result of domestication and human selection as proposed for grape (Matus et al., 2008). In *Citrus*, the presence of no more than one functional *MYB* regulator suggests that the production of anthocyanins is not under strong selective pressure and is largely dispensable. In Cucurbitaceae, the only family of plants where anthocyanin accumulation has never been reported, two *R2R3 MYB* genes of subgroup 6 are also present but, according to our analysis, neither gene encodes a functional MYB activator. Therefore, while in many plant species expansion of *MYB* genes has resulted in evolutionary novelty, in other lineages the loss of anthocyanin production is associated with molecular decay of the *MYB* regulators.

The *Ruby* locus is very dynamic. In a limited number of accessions, we uncovered six large deletions, six insertions of transposable elements, one translocation and four critical mutations within the *Ruby* gene and its promoter region (Fig. 3). Most of these events have had consequences on anthocyanin accumulation, ranging from a complete suppression of production to complex reconfiguration of the distribution of anthocyanins in different tissues.

Transposable elements are major forces in the evolution and rearrangement of genomes and important determinants of genome size and gene expression (McClintock, 1984). The insertion of Tc13 in citron likely did not result in obvious changes in pigmentation but the presence of Tcm7 in *C. micrantha*, accompanied by a deletion, was associated with almost complete abolition of anthocyanin

583 pigmentation in new growth.

584 The strong fruit-specific pigmentation of blood varieties of sweet orange was also the
585 result of the insertion of retrotransposons (Butelli et al., 2012). Comparative analysis
586 of core promoter regions in many different accessions and the mapping of the start
587 of transcription of *Ruby* in lemon allowed us to identify a key mutation in the TATA
588 box of *Ruby* in an allele derived from the pummelo parent in the hybrid, sweet
589 orange. As a result, sweet orange has completely lost the ability to accumulate
590 anthocyanins. In blood varieties of sweet orange, the insertion of retrotransposons
591 into this non-functional promoter prompted the use of alternative TATA boxes that
592 reestablished anthocyanin production and also determined the cold-induced
593 expression of *Ruby*, specifically in fruit (Butelli et al., 2012).

594 The accumulation of anthocyanins in fruit of species distantly related to *Citrus* (*M.*
595 *australasica* and *S. buxifolia*) was also associated with the insertion of transposable
596 elements upstream of *Ruby*. In this species, the retrotransposons belong to a
597 different family, SINE, and may have been involved in changing the activity of the
598 *Ruby* promoter and the control of anthocyanin production. Normally anthocyanin
599 production in plants is induced by light (Steyn et al., 2002) and the evolution of
600 anthocyanin production within the fruit might have required the development of light-
601 independent, regulatory control of *Ruby* expression, as observed with blood oranges
602 (Butelli et al., 2012)

603 The loss of anthocyanins may have occurred naturally, as a consequence of the
604 different climates in which *Citrus* species originated. Citron arose in high-altitude, dry
605 regions where anthocyanins are important in providing protection from high light and
606 UV irradiation (Moore, 1866; Steyn et al., 2002). In contrast, mandarin and pummelo
607 thrive in lowland tropics with high humidity and rain throughout the year (Webber,
608 1943). In these conditions, light stress may be less severe and anthocyanins less
609 important for protection of young tissues.

610 *Citrus* plants have been grown for millennia in Asia for reasons other than for food.
611 Citron has an ancient history as a sacred plant and its powerful symbolism is
612 associated with the fruit. Believed to be indigenous to Northeastern India, it was
613 absorbed into Jewish tradition and is still used in ancient religious rituals (Swingle
614 and Reece, 1967). China is the centre of origin of most citrus species. In Chinese
615 tradition, early use of *Citrus* was strictly religious and mainly associated with flowers.
616 Kumquat flowers represented luck and prosperity while the white flowers of mandarin

and orange were symbols of purity and innocence. This symbolism has persisted into modern Christianity (Heilmeyer, 2001; Moore and Garibaldi, 2003). The simultaneous occurrence of golden fruit and white blossom remains a central element in the symbolism of *C. tachibana*, a sacred tree in Japan. It is impossible to know the extent to which human intervention played a significant role in the selection of varieties of mandarin, pummelo and kumquat with pure white flowers for ornamental and religious purposes. Our study, however, indicates that the ability to accumulate anthocyanins has been lost independently, several times during *Citrus* evolution, as a result of mutations in the *Ruby* gene.

Analysis of the *Ruby* locus has informed us of the origin of present-day varieties of *Citrus* and the history of their spread around the world. Although phylogenetic inference based on the analysis of a single gene may lead to conclusions which require further support, our analysis of *Ruby* has confirmed the widely accepted hybrid origin of the most important cultivated *Citrus* varieties (Fig. 5) and the phylogenetic relationships between primary species and related genera (Fig. 7). The analysis of *Ruby* has also shed new light on the origin of poorly studied hybrids, on the classification of mandarins, on members of the diverse group *Papeda* and on the occurrence of ancient introgression events across species.

The analysis of *Ruby* explains why many *Citrus* varieties have lost the ability to accumulate anthocyanins and provides a useful tool to help in the interpretation of the 'mosaic' structure of *Citrus* genomes.

Materials and Methods

Plant Material

Leaves were obtained from different sources as indicated in Table 1. Plant materials were ground in liquid nitrogen and DNA and RNA were extracted using DNeasy Plant Mini Kit and RNeasy Plant Mini Kit (Qiagen) according to the manufacturer's instructions.

Isolation of *Ruby* alleles

For most accessions, *Ruby* alleles were isolated by PCR using primers designed to

amplify fragments spanning the region 1.6 kp upstream of the start of the coding sequence to 136 bp downstream of the *Ruby* stop codon. PCR fragments were used directly for sequencing or cloned into the pGEM-T Easy vector (Promega). For accessions where only direct sequencing was used, nucleotide polymorphisms were depicted using the IUPAC code. The active *Ruby* alleles in citron (R^C) and in *M. australasica* var 'Sanguinea' (R^{Mic2}) were isolated by inverse PCR after restriction enzyme digestion with *Pst*I and *Cla*I respectively followed by self-ligation. Primer sequences are listed in Supplemental Table S1.

DNA Gel Blot

DNA from leaves of lemons was digested with *Hind*III, separated by electrophoresis, transferred to nitrocellulose membranes and hybridized as described previously (Butelli et al., 2012).

Determination of the transcriptional start site of *Ruby*

The transcriptional start site in active alleles of *Ruby* was determined using RNA extracted from young purple leaves of 'Femminello' lemon. DNase-treated RNA was reverse transcribed using 5'/3' RACE Kit (Roche). The cDNA product was sequenced to determine the transcriptional start nucleotide in the sequences lying upstream of the *Ruby* coding sequences. Gene-specific primers listed in Supplemental Table S1 were used for 5' end cDNA amplification according to the manufacturer's instructions.

Expression analysis of *Ruby*

Quantification of *Ruby* expression in new growth of 'Femminello' and 'Zagara Bianca' lemons, 'Chandler' pummelo and in pigmented fruit and leaves of *M. australasica* var 'Sanguinea' was performed by qRT-PCR as previously described (Butelli et al., 2012). The allelic expression of *Ruby* in pigmented leaves of *P. trifoliata* and in fruit skin of *M. australasica* var 'Sanguinea' was determined by RT-PCR followed by sequencing.

Transient expression in *N. benthamiana* leaves

Transient expression of different *Ruby* cDNA constructs was carried out using the HyperTrans system (Sainsbury et al., 2009). *Ruby* cDNAs were isolated from the

flesh of 'Moro' blood orange or the fruit peel of *M. australasica* var 'Sanguinea', PCR amplified using primers listed in Supplemental Table S1 and transferred into pEAQ-HT-DEST1 vectors. A truncated cDNA containing the mutations identified in the alleles r^{STOPm} and r^{STOPp} was obtained using a reverse primer with a stop codon and cDNA from 'Moro' blood orange as template. The plasmids obtained were introduced into *Agrobacterium tumefaciens* GV3101 and the strains were infiltrated into the abaxial side of leaves of four-week-old plants. Leaves were harvested 7 days after infiltration.

Analysis and identification of anthocyanins

Leaves of lemon or *N. benthamiana* were frozen in liquid nitrogen and ground to a fine powder. Approximately 200 mg of powder was extracted with 800 μ L of 80% methanol containing 1% HCl. LC MS/MS of anthocyanins was carried out on a Nexera UHPLC system attached to an Ion-trap ToF (IT-ToF) mass spectrometer (Shimadzu). Separation was on a 100 \times 2.1mm 2.6 μ Kinetex XB-C18 column (Phenomenex) using the following gradient of acetonitrile (solvent B) versus 0.1% formic acid in water (solvent A), run at 0.5mL.min⁻¹ and 40°C: 0 min, 2% B; 0.5 min, 2% B; 2 min, 10% B; 11 min, 30% B; 15 min, 90% B; 15.8 min, 90% B; 16 min, 2% B; 20.1 min, 2% B. Anthocyanins were monitored by UV/visible absorbance (200-600nm) and positive mode electrospray MS (full spectra from m/z 200-2000 and data-dependent MS2 of the most abundant precursor ions at an isolation width of m/z 3.0, 50% collision energy and 50% collision gas). Anthocyanins were identified on the basis of their mass, and the mass of their fragments in MS2 as well as by comparison to published spectra.

Phylogenetic analysis

The sequences of the *Ruby* alleles were aligned using BioEdit (Hall, 1999), SeqMan version 7.0 (<http://www.dnastar.com>) and SATé-II (Liu et al., 2012) software applications. Analyses were performed to determine which model best matched the data using Phylemon 2.0 (<http://phylemon.bioinfo.cipf.es>; Sánchez et al., 2011). The HKY+G+F model took into account the nucleotide substitution model of Hasegawa-Kishino-Yano (HKY, nst=2), which allowed variable base frequencies, one transition rate and one transversion rate, and took into account the nucleotide frequency (F) and the gamma distribution (G). The construction of the maximum-likelihood (ML)

tree was performed using MEGA5 software with bootstrap support calculated from 1000 samples to assess the branch support.

Supplemental Data

The following supplemental materials are available.

Supplemental Figure S1. Classification of *Citrus* according to Swingle.

Supplemental Figure S2. Active Ruby alleles associated with LTR retrotransposons in blood varieties of sweet orange and citron.

Supplemental Figure S3. Analysis of promoter sequences in pummelo.

Supplemental Figure S4. Multiple sequence alignment of *Ruby* in pummelo.

Supplemental Figure S5. Multiple sequence alignment of *Ruby* in mandarin.

Supplemental Figure S6. Multiple sequence alignment of *Ruby* in different species of *Citrus*.

Supplemental Figure S7. PCR analysis of *Ruby* in primary species and hybrids of *Citrus*.

Supplemental Figure S8, S9 and S10. Mass spectrometry data for peak identification of anthocyanins in leaves of lemon.

Supplemental Figure S11. Expression analysis of *Ruby* in leaves of *P. trifoliata*.

Supplemental Figure S12. Expression analysis of *Ruby* in different tissues of *M. australasica* var 'Sanguinea'.

Supplemental Figure S13. Phylogenetic relationship between primary species of *Citrus* and related genera based on the sequence of *Ruby* using MEGA software program.

Supplemental Table S1. Sequences of the primers used in the study.

Supplemental Data Set S1. Nucleotide sequence of *Ruby2* in different accessions of the genus *Citrus*.

Figure and Table Legends

Table 1 *Citrus* accessions were obtained from the following sources: **A:** IVIA, Instituto Valenciano de Investigaciones Agrarias, Moncada, Valencia, Spain; **B:** CREA-CRA, Centro di Ricerca per l'Agrumicoltura e le Colture Mediterranee, Acireale, Italy; **C:** USDA-ARS National Clonal Germplasm Repository for Citrus & Dates, Riverside, CA, USA; **D:** CRB CITRUS, INRA-CIRAD, Citrus Biological Resource Center, San Giuliano, Corsica, France; **E:** Huazhong Agricultural University, Wuhan, China; **F:** Jingzhou, Huannan Province, China; **G:** Waitrose Supermarket, Norwich, UK; **H:** Notcutts Garden Centre, Norwich, UK.

In accessions of the genus *Citrus*, the *Ruby* alleles are colour-coded as follows: yellow, alleles derived from the mandarin pool; pink, alleles in mandarin introgressed from pummelo; red, alleles

derived from pummelo; green, alleles derived from citron; olive green, alleles derived from *Papeda*; the unique allele in *C. indica* is in blue; the alleles in genera related to *Citrus* are not coloured. The lemon accession indicated with an asterisk (*) is a 'graft chimera' and contains a third *Ruby* allele that may have originated from an acidless limetta. Pictures of fruit and flowers of the accessions indicated in bold are provided in Fig.1.

Figure 1. Phenotypes of some of the accessions considered in this study and listed in Table1. For each accession, the common and the scientific name according to Tanka's classification are provided, together with the allelic constitution at the *Ruby* locus. Images of fruit and flowers or young leaves are presented to show the presence or absence of anthocyanin pigmentation. Scale bar is approximately 1 cm. The images were obtained from the following sources: ♦, pictures taken by the authors; ☞, UC-Riverside Citrus Variety Collection (<http://www.citrusvariety.ucr.edu/index.html>); ☞, Citrus ID Tools (<http://idtools.org/id/citrus/citrusid/>); ⚭, Radoslav Krchnavek; ♂, Eve and George DeLange.

Figure 2. Functionality of different *Ruby* alleles. *Nicotiana benthamiana* leaves were infiltrated with cDNA constructs corresponding to three different *Ruby* alleles. The phenotypes of leaves, the colour of the leaf extracts and the HPLC chromatograms at 525 nm 7 days after infiltration are shown. A, Transient overexpression of *Ruby* cDNA encoding a full-length protein isolated from the flesh of 'Moro' blood orange. B, Transient overexpression of a cDNA construct encoding a protein with a terminal deletion of 116 amino acids corresponding to the alleles found in some pummelo and mandarin accessions and hybrids derived from them. C, Transient overexpression of *Ruby* cDNA isolated from the fruit peel of *Microcitrus australasica* var 'Sanguinea' encoding a protein with a 27 amino acid subterminal deletion. D, Evidence of the identity of delphinidin-3-rutinoside in sample C by positive mode electrospray MS1 and MS2. For this compound the expected mass is 611.1607 (found mass 611.1633; error 4.25ppm). Note the fragment representing delphinidin (expected mass 303.0499) and the losses of a hexose moiety (expected loss 162.0528) and rhamnose moiety (expected loss 146.0579).

Figure 3. Schematic representation of different *Ruby* alleles identified in primary species of *Citrus* and related genera. Functional alleles (*R*) associated with anthocyanin pigmentation are highlighted in purple.

Figure 4. Classification of mandarins based on the analysis of the *Ruby* gene. Mandarins can be classified into three groups (A, B and C) on the basis of two deletions and an introgression from pummelo. Many mandarin varieties are hybrids between members of these three groups or interspecific hybrids involving other *Citrus* species. *Citrus indica*, recognized by both Tanaka and Swingle as a mandarin species, is not related to mandarins. Roman numbers indicate the five taxonomic groups recognized by Tanaka (Hodgson, 1967). The gradation in colour shading indicates the contribution from genomes of other species. Common names for the mandarin accessions are provided in Table1.

Figure 5. Segregation analysis of *Ruby* in some of the commercially important *Citrus* species and hybrids considered in this study. Simplified diagram showing a single existing variety of pummelo and mandarin as potential parents of both sweet and sour orange, which is highly unlikely. Light purple boxes indicate species and hybrids that are able to accumulate anthocyanins and that are always associated with active *Ruby* alleles from citron (R^C) or papeda (R^{Pap2}). Pink and blue arrows indicate the maternal or parental contribution to the generation of a hybrid based on specific mitotypes and chlorotypes as determined in previously studies. The gradation in colour shading indicates the contribution from different species to the derivation of the hybrids.

Figure 6. Analysis of wild type and *white* mutant of lemon. A, HPLC chromatogram recorded at 525 nm of methanol extracts from young leaves of ‘Femminello’ wild type (purple line) and ‘Zagara Bianca’ (green line) lemons. Peaks marked with roman numbers indicate the main compounds identified in ‘Femminello’ and absent in ‘Zagara Bianca’. **Ia**, cyanidin 3-glucoside; **Ib**, cyanidin 3-rutinoside; **II**, cyanidin 3-(6"-malonyl)- β -glucoside; **III**, peonidin 3-(6"- malonylglucoside). Identification of peaks is based on UHPLC-ESI-MS analysis (Supplemental Fig. S8, S9 and S10) and comparison to literature data. B, Relative qRT-PCR quantification of *Ruby* expression in flower buds and young leaves of ‘Femminello’ and ‘Zagara Bianca’ lemons. Error bars represent SE of the mean. C, Schematic diagram of the two alleles of *Ruby* identified in lemon, R^C and r^{STOP} . The different symbols are explained in Fig. 3 and Supplemental Fig. S2A. The *HindIII* restriction sites and the genomic region used as a probe are shown. D, Southern blot analysis of genomic DNA from different lemon varieties digested with *HindIII* and probed with a 32 P-labelled probe of the *Ruby* gene. ‘Zagara Bianca’, the only variety completely unable to produce anthocyanins, doesn’t display the hybridization band corresponding to the active allele R^C .

Figure 7. Phylogenetic relationship between primary species of *Citrus* and related genera based on the sequence of *Ruby*. The phylogenetic trees were generated using Phylemon software program. Branch support is given in all branches. The different *Ruby* alleles are colour-coded as in Table 1. The sequences of the *Ruby* gene and its promoter used for phylogenetic analysis have been deposited in GenBank under accession numbers KT591672-KT591689.

References

- Barrett H, Rhodes AM** (1976) A numerical taxonomic study of affinity relationships in cultivated Citrus and its close relatives. *Sys Botany* **1**: 105–136
- Bayer RJ, Mabberley DJ, Morton C, Miller CH, Sharma IK, Pfeil BE, Rich S, Hitchcock R, Sykes S** (2009) A Molecular Phylogeny of the Orange Subfamily

- (Rutaceae: Aurantioideae) Using Nine Cpdna Sequences. American Journal of Botany **96**: 668-685
- Bombarely A, Moser M, Amrad A, Bao M, Bapaume L, Barry CS, Bliet M, Boersma MR, Borghi L, Bruggmann R et al.** (2016) Insight into the evolution of the Solanaceae from the parental genomes of *Petunia hybrida*. Nature Plants **2**: 16074
- Butelli E, Licciardello C, Zhang Y, Liu J, Mackay S, Bailey P, Reforgiato-Recupero G, Martin C** (2012) Retrotransposons control fruit-specific, cold-dependent accumulation of anthocyanins in blood oranges. Plant Cell **24**: 1242-1255
- Carbonell-Caballero J, Alonso R, Ibanez V, Terol J, Talon M, Dopazo J** (2015) A Phylogenetic Analysis of 34 Chloroplast Genomes Elucidates the Relationships between Wild and Domestic Species within the Genus *Citrus*. Molecular Biology and Evolution **32**: 2015-2035
- Coen ES, Carpenter R, Martin C** (1986) Transposable Elements Generate Novel Spatial Patterns of Gene-Expression in *Antirrhinum Majus*. Cell **47**: 285-296
- Cooper GM, Hausman RE** (2016) The Cell: A Molecular Approach. Sinauer Associates, Sunderland (MA)
- Cultrone A, Cotroneo PS, Recupero GR** (2010) Cloning and molecular characterization of R2R3-MYB and bHLH-MYC transcription factors from *Citrus sinensis*. Tree Genetics & Genomes **6**: 101-112
- Curk F, Ancillo G, Garcia-Lor A, Luro F, Perrier X, Jacquemoud-Collet JP, Navarro L, Ollitrault P** (2014) Next generation haplotyping to decipher nuclear genomic interspecific admixture in *Citrus* species: analysis of chromosome 2. BMC Genetics **15**: 152
- Curk F, Ancillo G, Ollitrault F, Perrier X, Jacquemoud-Collet JP, Garcia-Lor A, Navarro L, Ollitrault P** (2015) Nuclear Species-Diagnostic SNP Markers Mined from 454 Amplicon Sequencing Reveal Admixture Genomic Structure of Modern *Citrus* Varieties. Plos One **10**: e0125628
- Curk F, Ollitrault F, Garcia-Lor A, Luro F, Navarro L, Ollitrault P** (2016) Phylogenetic origin of limes and lemons revealed by cytoplasmic and nuclear markers. Annals of Botany **117**: 565-583
- Espley RV, Hellens RP, Putterill J, Stevenson DE, Kutty-Amma S, Allan AC** (2007) Red colouration in apple fruit is due to the activity of the MYB transcription factor, MdMYB10. Plant Journal **49**: 414-427
- Fabroni S, Ballistreri G, Amenta M, Rapisarda P** (2016) Anthocyanins in different *Citrus* species: an UHPLC-PDA-ESI/MSn-assisted qualitative and quantitative investigation. Journal of the Science of Food and Agriculture **96**: 4797-4808

- Federici CT, Fang DQ, Scora RW, Roose ML** (1998) Phylogenetic relationships within the genus *Citrus* (Rutaceae) and related genera as revealed by RFLP and RAPD analysis. *Theoretical and Applied Genetics* **96**: 812-822
- Fincham JRS** (1967) Mutable Genes in Light of Callans Hypothesis of Serially Repeated Gene Copies. *Nature* **215**: 864-866
- Froelicher Y, Mouhaya W, Bassene JB, Costantino G, Kamiri M, Luro F, Morillon R, Ollitrault P** (2011) New universal mitochondrial PCR markers reveal new information on maternal citrus phylogeny. *Tree Genetics & Genomes* **7**: 49-61
- Garcia-Lor A, Ancillo G, Navarro L, Ollitrault P** (2013) *Citrus* (Rutaceae) SNP Markers Based on Competitive Allele-Specific PCR; Transferability across the Aurantioideae Subfamily. *Applications in Plant Sciences* **1**: 1200406
- Garcia-Lor A, Curk F, Snoussi-Trifa H, Morillon R, Ancillo G, Luro F, Navarro L, Ollitrault P** (2013) A nuclear phylogenetic analysis: SNPs, indels and SSRs deliver new insights into the relationships in the 'true citrus fruit trees' group (Citrinae, Rutaceae) and the origin of cultivated species. *Annals of Botany* **111**: 1-19
- Garcia-Lor A, Luro F, Ollitrault P, Navarro L** (2015) Genetic diversity and population structure analysis of mandarin germplasm by nuclear, chloroplastic and mitochondrial markers. *Tree Genetics & Genomes* **11**: 123
- Hall T** (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* **41**: 95-98
- Heilmeyer M** (2001) *The Language of Flowers: Symbols and Myths*. Prestel USA, New York
- Hillebrand S, Schwarz M, Winterhalter P** (2004) Characterization of anthocyanins and pyranoanthocyanins from blood orange [*Citrus sinensis* (L.) Osbeck] juice. *Journal of Agricultural and Food Chemistry* **52**: 7331-7338
- Hoballah ME, Gubitz T, Stuurman J, Broger L, Barone M, Mandel T, Dell'Olio A, Arnold M, Kuhlemeier C** (2007) Single gene-mediated shift in pollinator attraction in *Petunia*. *Plant Cell* **19**: 779-790
- Hodgson RW** (1967) History, world distribution, botany and varieties. In W Reuther, H Webber, L Batchelor, eds, *The Citrus Industry*. University of California Press, Berkeley, CA, pp 431-591
- Kimura M** (1980) A Simple Method for Estimating Evolutionary Rates of Base Substitutions through Comparative Studies of Nucleotide-Sequences. *Journal of Molecular Evolution* **16**: 111-120
- Kitajima A, Yamasaki A, Habu T, Preedasuttijit B, Hasegawa K** (2007) Chromosome identification and karyotyping of satsuma mandarin by genomic in

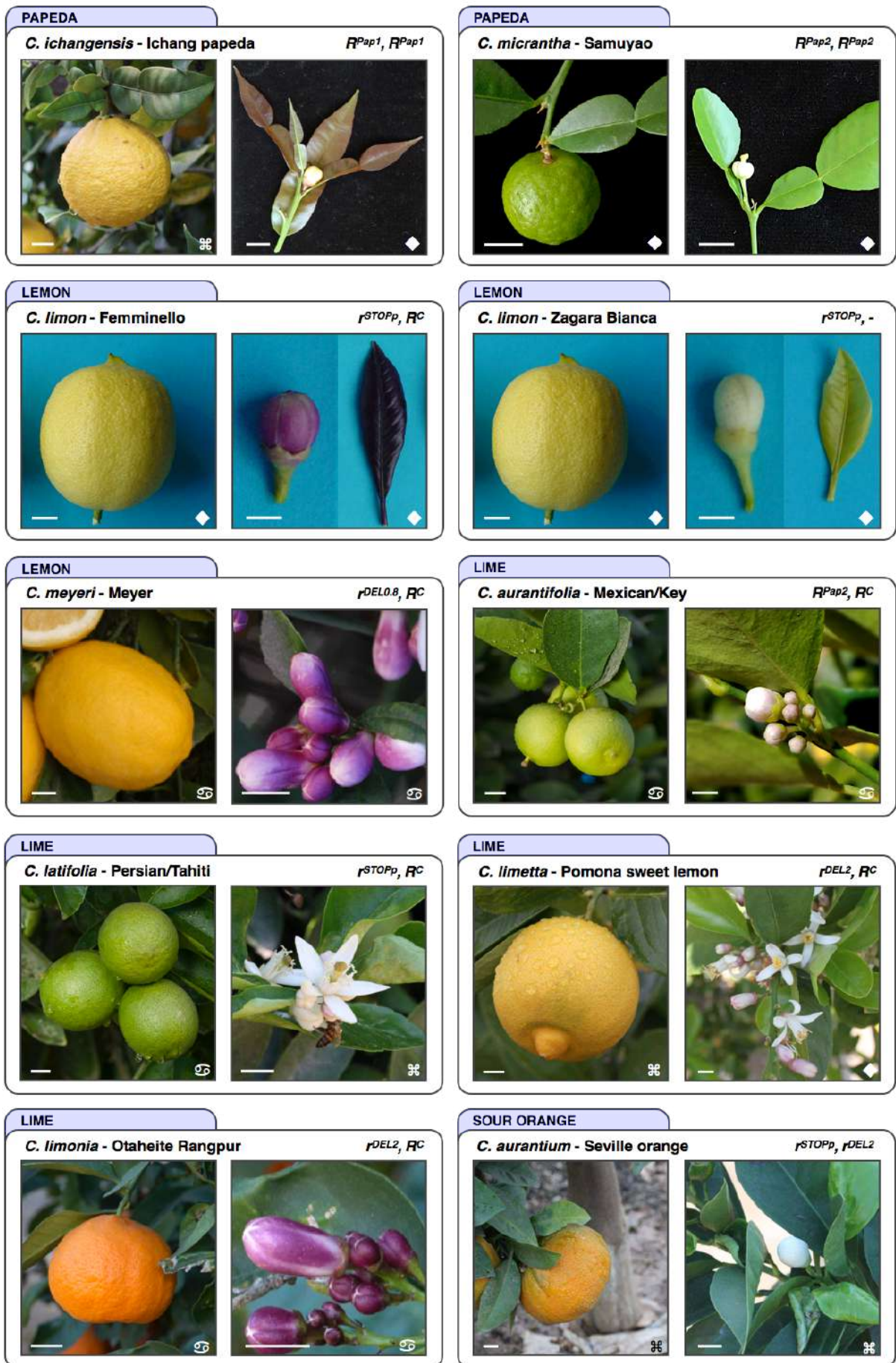
- situ hybridization. *Journal of the American Society for Horticultural Science* **132**: 836-841
- Kobayashi S, Goto-Yamamoto N, Hirochika H** (2004) Retrotransposon-induced mutations in grape skin color. *Science* **304**: 982
- Koes R, Verweij W, Quattrocchio F** (2005) Flavonoids: a colorful model for the regulation and evolution of biochemical pathways. *Trends in Plant Science* **10**: 236-242
- Krueger R, Navarro L** (2007) Citrus germplasm resources. *In* IA Khan, ed, Citrus genetics, breeding, and biotechnology. CAB International, Wallingford, UK, pp 45-140
- Lai YT, Chen IZ** (2008) Effect of Temperature on Calamondin (*Citrus microcarpa*) Flowering and Flower Bud Formation. *Proceedings of the International Symposium on Citrus and Other Tropical and Subtropical Fruit Crops*: 111-115
- Liu CY, Wang X, Xu YT, Deng XX, Xu Q** (2014) Genome-wide analysis of the R2R3-MYB transcription factor gene family in sweet orange (*Citrus sinensis*). *Molecular Biology Reports* **41**: 6769-6785
- Liu K, Warnow TJ, Holder MT, Nelesen SM, Yu JY, Stamatakis AP, Linder CR** (2012) SATE-II: Very Fast and Accurate Simultaneous Estimation of Multiple Sequence Alignments and Phylogenetic Trees. *Systematic Biology* **61**: 90-106
- Malinowski E** (1935) Studies on Unstable characters in *Petunia*. The extreme flower types of the Unstable race with mosaic color patterns. *Genetics* **20**: 342-356
- Matus JT, Aquea F, Arce-Johnson P** (2008) Analysis of the grape MYB R2R3 subfamily reveals expanded wine quality-related clades and conserved gene structure organization across *Vitis* and *Arabidopsis* genomes. *Bmc Plant Biology* **8**: 83
- McClintock B** (1984) The Significance of Responses of the Genome to Challenge. *Science* **226**: 792-801
- Moore A, Garibaldi C** (2003) Flower Power: The Meaning of Flowers in Art 1500-2000. Philip Wilson Publishers, London
- Moore GA** (2001) Oranges and lemons: clues to the taxonomy of Citrus from molecular markers. *Trends in Genetics* **17**: 536-540
- Moore T** (1866) The Treasury of Botany: a popular dictionary of the vegetable kingdom. , Vol 1. Longmans Green & Co., London
- Morton JF** (1987) Grapefruit. *In* Fruits of warm climates. University of Florida, Miami, pp 152-158

- Netzel M, Netzel G, Tian QG, Schwartz S, Konczak I** (2006) Sources of antioxidant activity in Australian native fruits. Identification and quantification of anthocyanins. *Journal of Agricultural and Food Chemistry* **54**: 9820-9826
- Nicolosi E, Deng ZN, Gentile A, La Malfa S, Continella G, Tribulato E** (2000) Citrus phylogeny and genetic origin of important species as investigated by molecular markers. *Theoretical and Applied Genetics* **100**: 1155-1166
- Penjor T, Anai T, Nagano Y, Matsumoto R, Yamamoto M** (2010) Phylogenetic relationships of Citrus and its relatives based on rbcL gene sequences. *Tree Genetics & Genomes* **6**: 931-939
- Penjor T, Yamamoto M, Uehara M, Ide M, Matsumoto N, Matsumoto R, Nagano Y** (2013) Phylogenetic Relationships of Citrus and Its Relatives Based on matK Gene Sequences. *Plos One* **8**: e62574
- Quattrocchio F, Wing JF, Leppen HTC, Mol JNM, Koes RE** (1993) Regulatory Genes-Controlling Anthocyanin Pigmentation Are Functionally Conserved among Plant-Species and Have Distinct Sets of Target Genes. *Plant Cell* **5**: 1497-1512
- Quattrocchio F, Wing J, van der Woude K, Souer E, de Vetten N, Mol J, Koes R** (1999) Molecular analysis of the anthocyanin2 gene of petunia and its role in the evolution of flower color. *Plant Cell* **11**: 1433-1444
- Ramadugu C, Pfeil BE, Keremane ML, Lee RF, Maureira-Butler IJ, Roose ML** (2013) A Six Nuclear Gene Phylogeny of Citrus (Rutaceae) Taking into Account Hybridization and Lineage Sorting. *Plos One* **8**: e68410
- Ramsay NA, Glover BJ** (2005) MYB-bHLH-WD40 protein complex and the evolution of cellular diversity. *Trends in Plant Science* **10**: 63-70
- Sainsbury F, Thuenemann EC, Lomonossoff GP** (2009) pEAQ: versatile expression vectors for easy and quick transient expression of heterologous proteins in plants. *Plant Biotechnology Journal* **7**: 682-693
- Sanchez R, Serra F, Tarraga J, Medina I, Carbonell J, Pulido L, de Maria A, Capella-Gutierrez S, Huerta-Cepas J, Gabaldon T, Dopazo J, Dopazo H** (2011) Phylemon 2.0: a suite of web-tools for molecular evolution, phylogenetics, phylogenomics and hypotheses testing. *Nucleic Acids Research* **39**: W470-W474
- SanMiguel P, Gaut BS, Tikhonov A, Nakajima Y, Bennetzen JL** (1998) The paleontology of intergene retrotransposons of maize. *Nature Genetics* **20**: 43-45
- Schwinn K, Venail J, Shang YJ, Mackay S, Alm V, Butelli E, Oyama R, Bailey P, Davies K, Martin C** (2006) A small family of MYB-regulatory genes controls floral pigmentation intensity and patterning in the genus *Antirrhinum*. *Plant Cell* **18**: 831-851

- Scora RW** (1975) Symposium on Biochemical Systematics, Genetics and Origin of Cultivated Plants .9. History and Origin of Citrus. Bulletin of the Torrey Botanical Club **102**: 369-375
- Steyn WJ, Wand SJE, Holcroft DM, Jacobs G** (2002) Anthocyanins in vegetative tissues: a proposed unified function in photoprotection. New Phytologist **155**: 349-361
- Stracke R, Werber M, Weisshaar B** (2001) The R2R3-MYB gene family in *Arabidopsis thaliana*. Current Opinion in Plant Biology **4**: 447-456
- Swingle WT** (1946) The botany of Citrus and its wild relatives of the orange subfamily (family Rutaceae, subfamily Aurantioideae). *In* W Reuther, H Webber, L Batchelor, eds, The Citrus Industry. University of California Press, Berkeley, CA, pp 129-474
- Swingle WT, Reece PC** (1967) The botany of Citrus and its wild relatives. *In* W Reuther, H Webber, L Batchelor, eds, The Citrus Industry. University of California Press, Berkeley, CA, pp 190–430
- Tamura K** (1992) The rate and pattern of nucleotide substitution in *Drosophila* mitochondrial DNA. Mol Biol Evol **9**: 814-825
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S** (2011) MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Molecular Biology and Evolution **28**: 2731-2739
- Tanaka T** (1977) Fundamental discussion of citrus classification. Studia Citrologica **14**: 1-6
- Tanaka Y, Sasaki N, Ohmiya A** (2008) Biosynthesis of plant pigments: anthocyanins, betalains and carotenoids. Plant Journal **54**: 733-749
- Webber HJ** (1943) Plant characteristics and climatology. *In* W Reuther, H Webber, L Batchelor, eds, The Citrus Industry, Vol 1. University of California Press, Berkeley, CA, pp 41-69
- Winkel-Shirley B** (2001) Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. Plant Physiology **126**: 485-493
- Wu GA, Prochnik S, Jenkins J, Salse J, Hellsten U, Murat F, Perrier X, Ruiz M, Scalabrin S, Terol J et al.** (2014) Sequencing of diverse mandarin, pummelo and orange genomes reveals complex history of admixture during citrus domestication. Nature Biotechnology **32**: 656-662
- Xu Q, Chen LL, Ruan XA, Chen DJ, Zhu AD, Chen CL, Bertrand D, Jiao WB, Hao BH, Lyon MP et al.** (2013) The draft genome of sweet orange (*Citrus sinensis*). Nature Genetics **45**: 59-U92

1086
1087
1088
1089
1090
1091





GRAPEFRUIT

C. paradisi - Marsh

r^{STOPp} , r^{DEL2}



SWEET ORANGE

C. sinensis - Navel

r^{TATA} , r^{DEL2}



SWEET ORANGE

C. sinensis - Moro

$R^{D.2}$, r^{DEL2}



PONCIRUS

P. trifoliata - trifoliate orange

R^{Fon} , R^{Pon}



MICROCITRUS

M. australasica - Green finger lime

R^{Mic1} , R^{Mic1}



MICROCITRUS

M. australasica - Sanguinea finger lime

R^{Mic1} , R^{Mic2}



FORTUNELLA

F. margarita - Nagami oval kumquat

r^{For} , r^{For}



FORTUNELLA

F. polyandra - Malayan kumquat

r^{For} , r^{For}



FORTUNELLA

F. madurensis - Calamondin

r^{For} , $r^{DEL0.8}$



SEVERINIA

S. buxifolia - Chinese box orange

R^{Sev} , R^{Sev}



Figure 1. Phenotypes of some of the accessions considered in this study and listed in Table 1. For each accession, the common and the scientific name according to Tanka's classification are provided, together with the allelic constitution at the *Ruby* locus. Images of fruit and flowers or young leaves are presented to show the presence or absence of anthocyanin pigmentation. Scale bar is approximately 1 cm. The images were obtained from the following sources: ♦, pictures taken by the authors; ☞, UC-Riverside Citrus Variety Collection (<http://www.citrusvariety.ucr.edu/index.html>); ⌘, Citrus ID Tools (<http://idtools.org/id/citrus/citrusid/>); ⤵, Radoslav Krchnavek; ⚡; Eve and George DeLange.

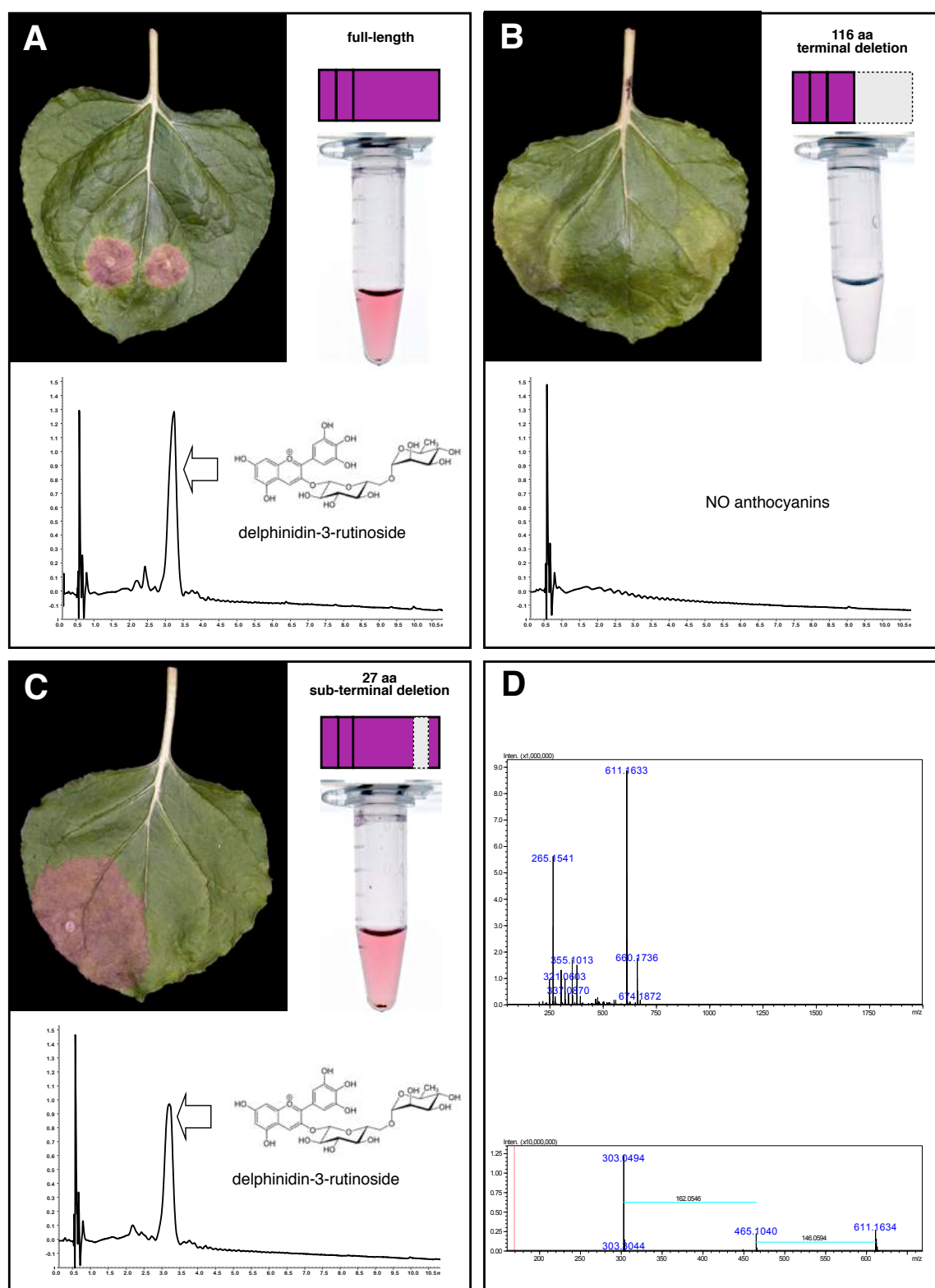


Figure 2.

Figure 2. Functionality of different *Ruby* alleles. *Nicotiana benthamiana* leaves were infiltrated with cDNA constructs corresponding to three different *Ruby* alleles. The phenotypes of leaves, the colour of the leaf extracts and the HPLC chromatograms at 525 nm 7 days after infiltration are shown. A, Transient overexpression of *Ruby* cDNA encoding a full-length protein isolated from the flesh of ‘Moro’ blood orange. B, Transient overexpression of a cDNA construct encoding a protein with a terminal deletion of 116 amino acids corresponding to the alleles found in some pummelo and mandarin accessions and hybrids derived from them. C, Transient overexpression of *Ruby* cDNA isolated from the fruit peel of *Microcitrus australasica* var ‘Sanguinea’ encoding a protein with a 27 amino acid subterminal deletion. D, Evidence of the identity of delphinidin-3-rutinoside in sample C by positive mode electrospray MS1 and MS2. For this compound the expected mass is 611.1607 (found mass 611.1633; error 4.25ppm). Note the fragment representing delphinidin (expected mass 303.0499) and the losses of a hexose moiety (expected loss 162.0528) and rhamnose moiety (expected loss 146.0579).

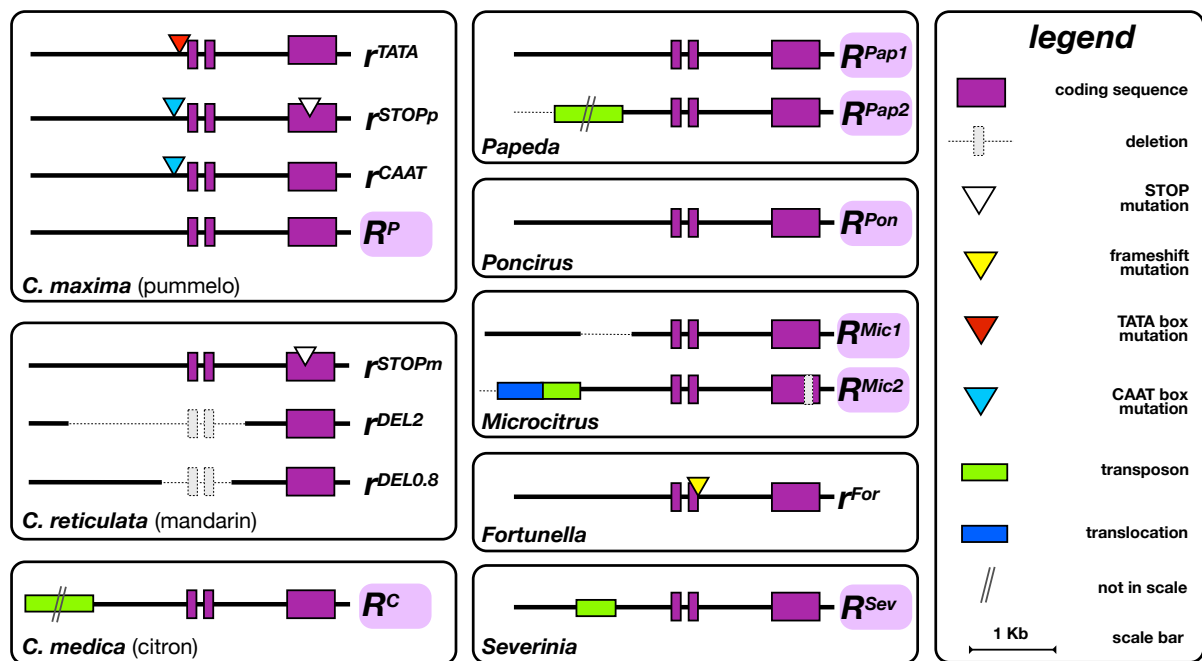


Figure 3. Schematic representation of different *Ruby* alleles identified in primary species of *Citrus* and related genera. Functional alleles (*R*) associated with anthocyanin pigmentation are highlighted in purple.

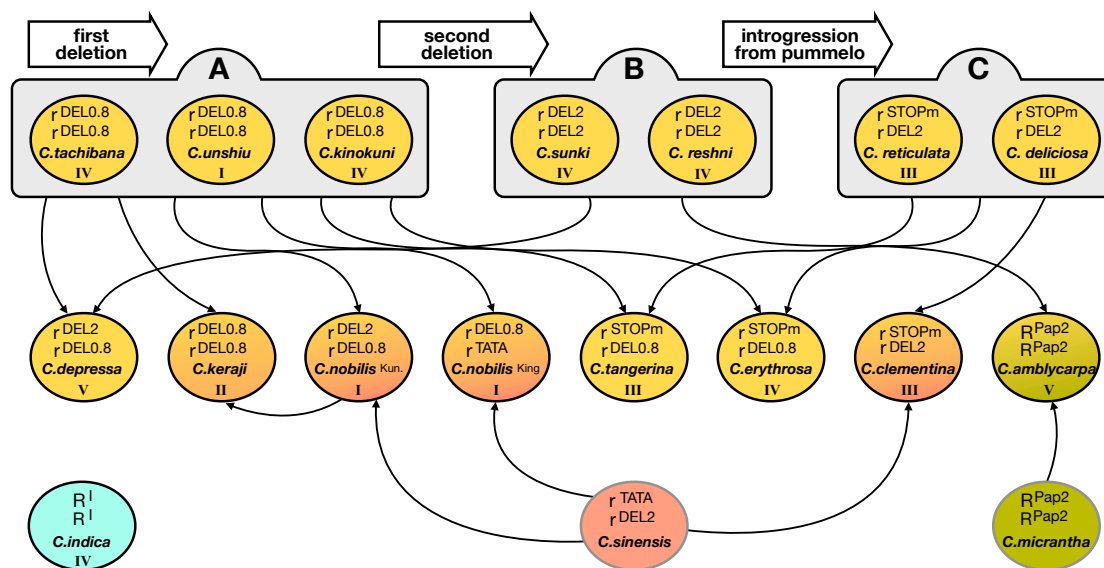


Figure 4. Classification of mandarins based on the analysis of the *Ruby* gene. Mandarins can be classified into three groups (A, B and C) on the basis of two deletions and an introgression from pummelo. Many mandarin varieties are hybrids between members of these three groups or interspecific hybrids involving other *Citrus* species. *Citrus indica*, recognized by both Tanaka and Swingle as a mandarin species, is not related to mandarins. Roman numbers indicate the five taxonomic groups recognized by Tanaka (Hodgson, 1967). The gradation in colour shading indicates the contribution from genomes of other species. Common names for the mandarin accessions are provided in Table1.

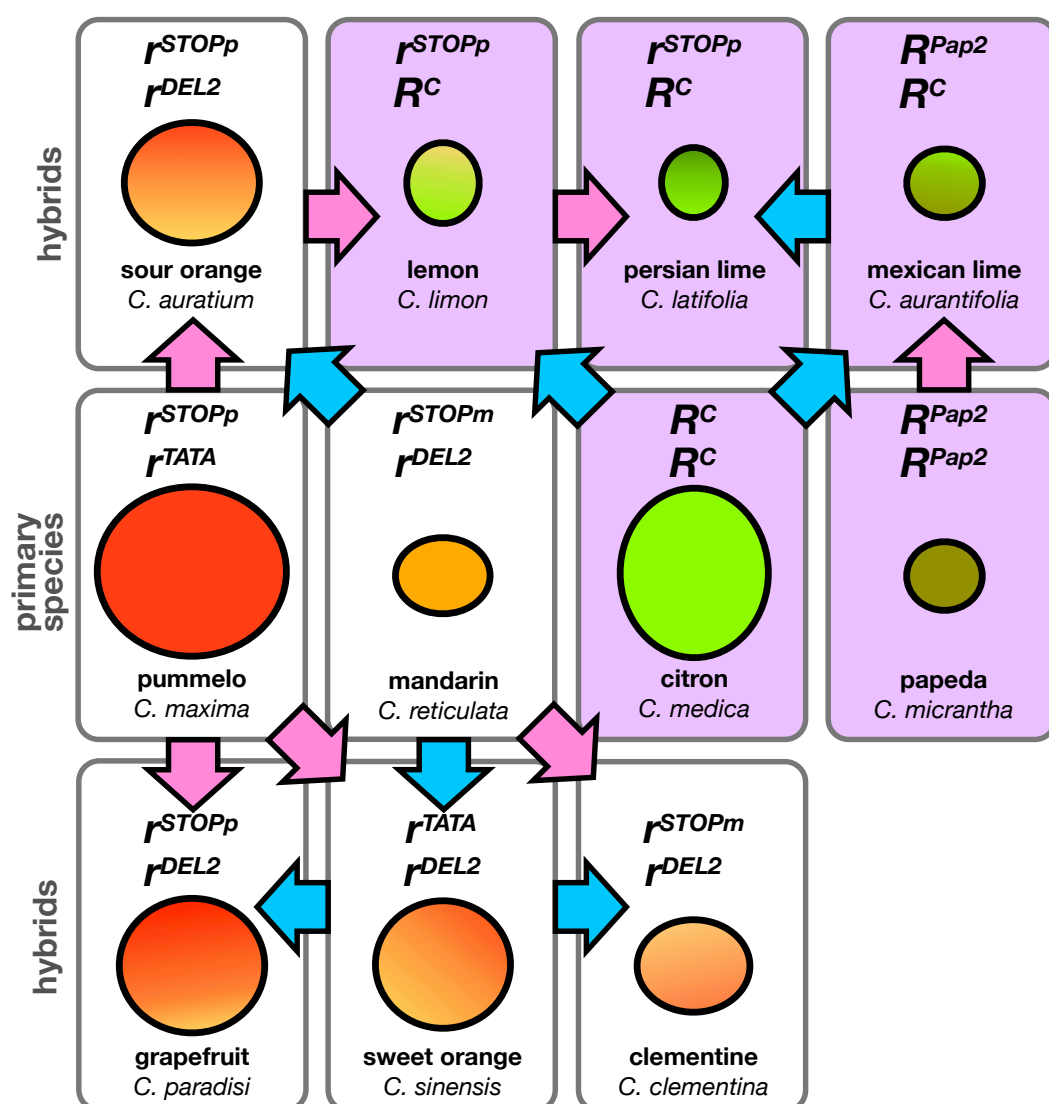


Figure 5. Segregation analysis of *Ruby* in some of the commercially important *Citrus* species and hybrids considered in this study. Simplified diagram showing a single existing variety of pummelo and mandarin as potential parents of both sweet and sour orange, which is highly unlikely. Light purple boxes indicate species and hybrids that are able to accumulate anthocyanins and that are always associated with active *Ruby* alleles from citron (R^C) or papeda (R^{Pap2}). Pink and blue arrows indicate the maternal or parental contribution to the generation of a hybrid based on specific mitotypes and chlorotypes as determined in previously studies. The gradation in colour shading indicates the contribution from different species to the derivation of the hybrids.

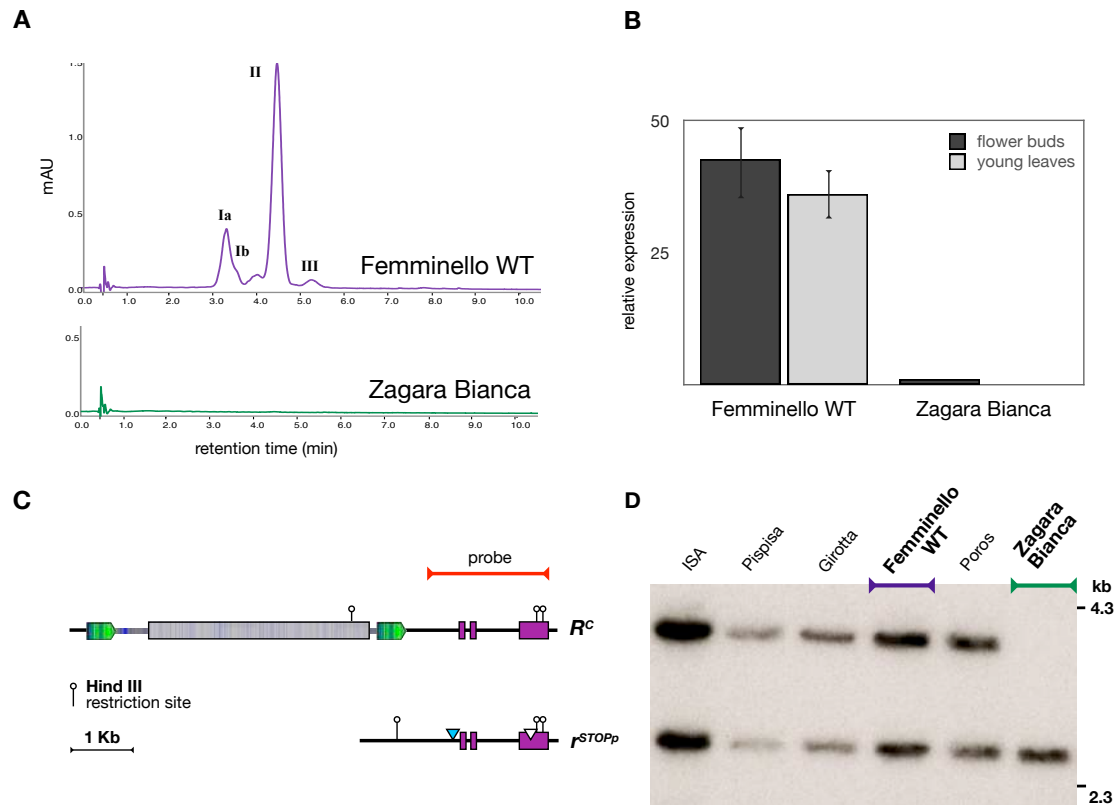


Figure 6. Analysis of wild type and *white* mutant of lemon. **A**, HPLC chromatogram recorded at 525 nm of methanol extracts from young leaves of ‘Femminello’ wild type (purple line) and ‘Zagara Bianca’ (green line) lemons. Peaks marked with roman numbers indicate the main compounds identified in ‘Femminello’ and absent in ‘Zagara Bianca’. **Ia**, cyanidin 3-glucoside; **Ib**, cyanidin 3-rutinoside; **II**, cyanidin 3-(6"-malonyl)- β -glucoside; **III**, peonidin 3-(6"-malonylglucoside). Identification of peaks is based on UHPLC-ESI-MS analysis (Supplemental Fig. S8, S9 and S10) and comparison to literature data. **B**, Relative qRT-PCR quantification of *Ruby* expression in flower buds and young leaves of ‘Femminello’ and ‘Zagara Bianca’ lemons. Error bars represent SE of the mean. **C**, Schematic diagram of the two alleles of *Ruby* identified in lemon, R^C and r^{STOPp} . The different symbols are explained in Fig. 3 and Supplemental Fig. S2A. The *Hind*III restriction sites and the genomic region used as a probe are shown. **D**, Southern blot analysis of genomic DNA from different lemon varieties digested with *Hind*III and probed with a 32 P-labelled probe of the *Ruby* gene. ‘Zagara Bianca’, the only variety completely unable to produce anthocyanins, doesn’t display the hybridization band corresponding to the active allele R^C .

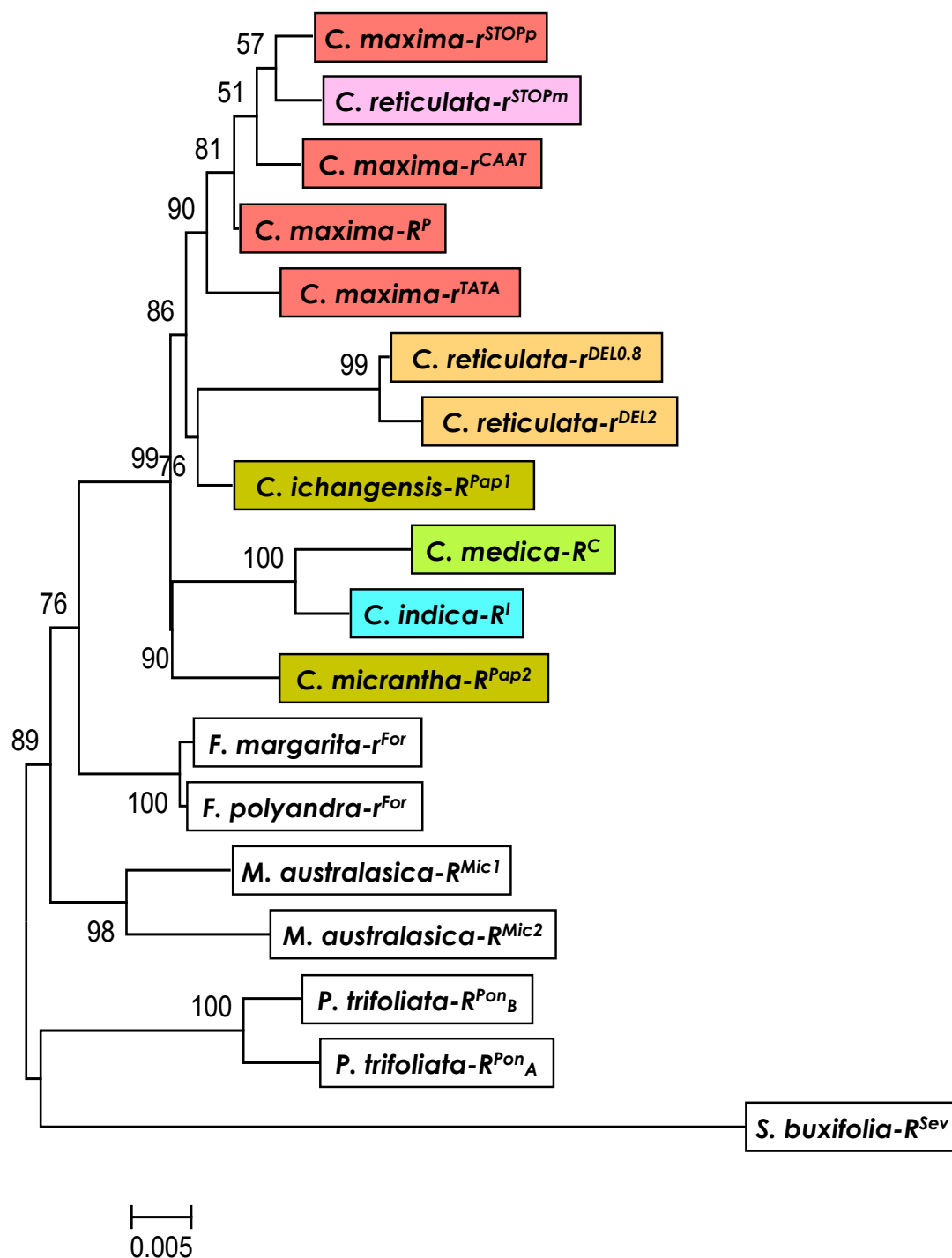


Figure 7. Phylogenetic relationship between primary species of *Citrus* and related genera based on the sequence of *Ruby*. The phylogenetic trees were generated using Phylemon software program. Branch support is given in all branches. The different *Ruby* alleles are colour-coded as in Table 1. The sequences of the *Ruby* gene and its promoter used for phylogenetic analysis have been deposited in GenBank under accession numbers KT591672-KT591689.

Table 1. Species and hybrids used in this study

Common name	Swingle	Tanaka	Ruby alleles		Source	Accession
CITRON						
Diamante	<i>C. medica</i>	<i>C. medica</i>	<i>R^C</i>	<i>R^C</i>	B	Palazzelli Certif.
Corsican	<i>C. medica</i>	<i>C. medica</i>	<i>R^C</i>	<i>R^C</i>	D	SRA613
Poncire commun	<i>C. medica</i>	<i>C. medica</i>	<i>R^C</i>	<i>R^C</i>	D	SRA701
Buddha's Hand	<i>C. medica</i>	<i>C. medica</i>	<i>R^C</i>	<i>R^C</i>	D	SRA 640
PUMMELO						
Chandler	<i>C. maxima</i>	<i>C. maxima</i>	<i>r^{TATA}</i>	<i>r^{CAAT}</i>	B	1A-F7P13
Reinking	<i>C. maxima</i>	<i>C. maxima</i>	<i>r^{TATA}</i>	<i>r^{STOPp}</i>	B	1A-F9P16
Purple-skin	<i>C. maxima</i>	<i>C. maxima</i>	<i>r^{TATA}</i>	<i>R^P</i>	E	-
Kao Pan	<i>C. maxima</i>	<i>C. maxima</i>	<i>r^{CAAT}</i>	<i>r^{STOPp}</i>	E	-
Siamese acidless	<i>C. maxima</i>	<i>C. maxima</i>	<i>r^{CAAT}</i>	<i>r^{CAAT}</i>	E	-
MANDARIN						
Satsuma Frost Owari	<i>C. reticulata</i>	<i>C. unshiu</i>	<i>r^{DEL0.8}</i>	<i>r^{DEL0.8}</i>	A	IVIA175
Satsuma Okitsu	<i>C. reticulata</i>	<i>C. unshiu</i>	<i>r^{DEL0.8}</i>	<i>r^{DEL0.8}</i>	A	IVIA195
Kishu	<i>C. reticulata</i>	<i>C. kinokuni</i>	<i>r^{DEL0.8}</i>	<i>r^{DEL0.8}</i>	A	IVIA678
Sunki	<i>C. reticulata</i>	<i>C. sunki</i>	<i>r^{DEL2}</i>	<i>r^{DEL2}</i>	A	IVIA239
Cleopatra	<i>C. reticulata</i>	<i>C. reshni</i>	<i>r^{DEL2}</i>	<i>r^{DEL2}</i>	A	IVIA385
Ponkan	<i>C. reticulata</i>	<i>C. reticulata</i>	<i>r^{DEL2}</i>	<i>r^{STOPm}</i>	A,B	IVIA482; A-F10P2
Avana	<i>C. reticulata</i>	<i>C. deliciosa</i>	<i>r^{DEL2}</i>	<i>r^{STOPm}</i>	A,B	IVIA189; A-F21P3
Dancy	<i>C. reticulata</i>	<i>C. tangerina</i>	<i>r^{DEL0.8}</i>	<i>r^{STOPm}</i>	A	IVIA434
Fuzhu	<i>C. reticulata</i>	<i>C. erythrosa</i>	<i>r^{DEL0.8}</i>	<i>r^{STOPm}</i>	A	IVIA571
Shekwasha	<i>C. reticulata</i>	<i>C. depressa</i>	<i>r^{DEL0.8}</i>	<i>r^{DEL2}</i>	A	IVIA238
King	<i>C. reticulata</i>	<i>C. nobilis</i>	<i>r^{DEL0.8}</i>	<i>r^{TATA}</i>	A	IVIA477
Kunembo	<i>C. reticulata</i>	<i>C. nobilis</i>	<i>r^{DEL2}</i>	<i>r^{DEL0.8}</i>	C	CRC3346
Keraji	<i>C. reticulata</i>	<i>C. keraji</i>	<i>r^{DEL0.8}</i>	<i>r^{DEL0.8}</i>	C	CRC3144
Oroval clementine	<i>C. reticulata</i>	<i>C. clementina</i>	<i>r^{DEL2}</i>	<i>r^{STOPm}</i>	B	Acireale Certif.
Nasranan	<i>C. reticulata</i>	<i>C. amblycarpa</i>	<i>R^{Pap2}</i>	<i>R^{Pap2}</i>	A	IVIA478
Tachibana orange	<i>C. tachibana</i>	<i>C. tachibana</i>	<i>r^{DEL0.8}</i>	<i>r^{DEL0.8}</i>	A,C	IVIA237; CRC3150
Indian wild orange	<i>C. indica</i>	<i>C. indica</i>	<i>R^C</i>	<i>R^C</i>	A,C	IVIA550; CRC3163
PAPEDA						
Kaffir lime	<i>C. hystrix</i>	<i>C. hystrix</i>	<i>R^{Pap1}</i>	<i>R^{Pap2}</i>	A	IVIA178
Melanesian papeda	<i>C. macroptera</i>	<i>C. macroptera</i>	<i>R^{Pap1}</i>	<i>R^{Pap2}</i>	A	IVIA279
Ichang papeda	<i>C. ichangensis</i>	<i>C. ichangensis</i>	<i>R^{Pap1}</i>	<i>R^{Pap1}</i>	A	IVIA235
Samuyao	<i>C. micrantha</i> var <i>microcarpa</i>	<i>C. micrantha</i>	<i>R^{Pap2}</i>	<i>R^{Pap2}</i>	A	IVIA626
LEMON						
Femminello	<i>C. limon</i>	<i>C. limon</i>	<i>r^{STOPp}</i>	<i>R^C</i>	B	C-F8P2
Politi	<i>C. limon</i>	<i>C. limon</i>	<i>r^{STOPp}</i>	<i>R^C</i>	B	C-F3P1
Zagara Bianca	<i>C. limon</i>	<i>C. limon</i>	<i>r^{STOPp}</i>	-	B	B-F7P3
Dulce*	<i>C. limon</i>	<i>C. limon</i>	<i>r^{STOPp}</i>	<i>R^C</i>	H	IVIA443
Meyer	<i>C. limon</i> hybrid	<i>C. meyeri</i>	<i>r^{DEL0.8}</i>	<i>R^C</i>	H	-
LIME						
Mexican/Key	<i>C. aurantifolia</i>	<i>C. aurantifolia</i>	<i>R^{Pap2}</i>	<i>R^C</i>	A	IVIA164
Persian/Tahiti	<i>C. aurantifolia</i>	<i>C. latifolia</i>	<i>r^{STOPp}</i>	<i>R^C</i>	G	-
Palestine sweet lime	<i>C. aurantifolia</i>	<i>C. limettioides</i>	<i>r^{DEL0.8}</i>	<i>R^C</i>	C	CRC1482
Mary Ellen	<i>C. aurantifolia</i>	<i>C. limettioides</i>	<i>r^{DEL0.8}</i>	<i>R^C</i>	C	CRC4053
Soh Synteng	<i>C. limon</i>	<i>C. limon</i>	<i>r^{DEL0.8}</i>	<i>R^C</i>	C	CRC3261
Limonette de Marrakech	<i>C. limon</i>	<i>C. limetta</i>	<i>r^{DEL2}</i>	<i>R^C</i>	C,D	CRC3989; SRA829
Limetta Dolce	<i>C. limon</i>	<i>C. limetta</i>	<i>r^{DEL2}</i>	<i>R^C</i>	B	C-F2P7
Pomona sweet lemon	<i>C. limon</i>	<i>C. limetta</i>	<i>r^{DEL2}</i>	<i>R^C</i>	C	CRC4068
Rangpur Kirumakki	<i>C. limon</i>	<i>C. limonia</i>	<i>r^{DEL2}</i>	<i>R^C</i>	C	CRC4131
Rangpur Otaheite	<i>C. limon</i>	<i>C. limonia</i>	<i>r^{DEL2}</i>	<i>R^C</i>	C	CRC2709
SOUR ORANGE						
Seville orange	<i>C. aurantium</i>	<i>C. aurantium</i>	<i>r^{STOPp}</i>	<i>r^{DEL2}</i>	G	-
GRAPEFRUIT						
Marsh	<i>C. paradisi</i>	<i>C. paradisi</i>	<i>r^{STOPp}</i>	<i>r^{DEL2}</i>	G	-
SWEET ORANGE						
Navel	<i>C. sinensis</i>	<i>C. sinensis</i>	<i>r^{TATA}</i>	<i>r^{DEL2}</i>	B	2B-F1-P14
Valencia	<i>C. sinensis</i>	<i>C. sinensis</i>	<i>r^{TATA}</i>	<i>r^{DEL2}</i>	B	6A2-F5P1
Tarocco blood	<i>C. sinensis</i>	<i>C. sinensis</i>	<i>R^{D.1}</i>	<i>r^{DEL2}</i>	B	4-F20P9
Moro blood	<i>C. sinensis</i>	<i>C. sinensis</i>	<i>R^{D.2}</i>	<i>r^{DEL2}</i>	B	8-F1P2
Jingxian blood	<i>C. sinensis</i>	<i>C. sinensis</i>	<i>R^{D.3}</i>	<i>r^{DEL2}</i>	F	-

Citrus-Related Genera						
PONCIRUS						
Trifoliolate orange	<i>P. trifoliata</i>	<i>P. trifoliata</i>	R^{Pon}	R^{Pon}	B	-
MICROCITRUS						
Australian finger lime	<i>M. australasica</i> 'green fruit'	<i>M. australasica</i>	R^{Mic1}	R^{Mic1}	C	CRC3661
Australian finger lime	<i>M. australasica</i> var <i>Sanguinea</i>	<i>M. australasica</i>	R^{Mic2}	R^{Mic1}	B	C-F5P3
FORTUNELLA						
Nagami oval kumquat	<i>F. margarita</i>	<i>F. margarita</i>	r^{For}	r^{For}	H	-
Malayan kumquat	<i>F. polyandra</i>	<i>F. polyandra</i>	r^{For}	r^{For}	A	IVIA375
Calamondin	<i>Citrofortunella</i> <i>microcarpa</i>	<i>C. madurensis</i>	r^{For}	$r^{DEL0.8}$	H	-
SEVERINIA						
Chinese box orange	<i>S. buxifolia</i>	<i>S. buxifolia</i>	R^{Sev}	R^{Sev}	A	IVIA147

Citrus accessions were obtained from the following sources: **A**: IVIA, Instituto Valenciano de Investigaciones Agrarias, Moncada, Valencia, Spain; **B**: CREA-CRA, Centro di Ricerca per l'Agricoltura e le Colture Mediterranee, Acireale, Italy; **C**: USDA-ARS National Clonal Germplasm Repository for Citrus & Dates, Riverside, CA, USA; **D**: CRB CITRUS, INRA-CIRAD, Citrus Biological Resource Center, San Giuliano, Corsica, France; **E**: Huazhong Agricultural University, Wuhan, China; **F**: Jingzhou, Huannan Province, China; **G**: Waitrose Supermarket, Norwich, UK; **H**: Notcutts Garden Centre, Norwich, UK.

In accessions of the genus *Citrus*, the *Ruby* alleles are colour-coded as follows: yellow, alleles derived from the mandarin pool; pink, alleles in mandarin introgressed from pummelo; red, alleles derived from pummelo; green, alleles derived from citron; olive green, alleles derived from *Papeda*; the unique allele in *C. indica* is in blue; the alleles in genera related to *Citrus* are not coloured. The lemon accession indicated with an asterisk (*) is a 'graft chimera' and contains a third *Ruby* allele that may have originated from an acidless limetta. Pictures of fruit and flowers of the accessions indicated in bold are provided in Fig.1.

Parsed Citations

Barrett H, Rhodes AM (1976) A numerical taxonomic study of affinity relationships in cultivated Citrus and its close relatives. Sys Botany 1: 105-136

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Bayer RJ, Mabberley DJ, Morton C, Miller CH, Sharma IK, Pfeil BE, Rich S, Hitchcock R, Sykes S (2009) A Molecular Phylogeny of the Orange Subfamily (Rutaceae: Aurantioideae) Using Nine Cpdna Sequences. American Journal of Botany 96: 668-685

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Bombarely A, Moser M, Amrad A, Bao M, Bapaume L, Barry CS, Bliek M, Boersma MR, Borghi L, Bruggmann R et al. (2016) Insight into the evolution of the Solanaceae from the parental genomes of Petunia hybrida. Nature Plants 2: 16074

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Butelli E, Licciardello C, Zhang Y, Liu J, Mackay S, Bailey P, Reforgiato-Recupero G, Martin C (2012) Retrotransposons control fruit-specific, cold-dependent accumulation of anthocyanins in blood oranges. Plant Cell 24: 1242-1255

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Carbonell-Caballero J, Alonso R, Ibanez V, Terol J, Talon M, Dopazo J (2015) A Phylogenetic Analysis of 34 Chloroplast Genomes Elucidates the Relationships between Wild and Domestic Species within the Genus Citrus. Molecular Biology and Evolution 32: 2015-2035

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Coen ES, Carpenter R, Martin C (1986) Transposable Elements Generate Novel Spatial Patterns of Gene-Expression in Antirrhinum Majus. Cell 47: 285-296

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Cooper GM, Hausman RE (2016) The Cell: A Molecular Approach. Sinauer Associates, Sunderland (MA)

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Cultrone A, Cotroneo PS, Recupero GR (2010) Cloning and molecular characterization of R2R3-MYB and bHLH-MYC transcription factors from Citrus sinensis. Tree Genetics & Genomes 6: 101-112

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Curk F, Ancillo G, Garcia-Lor A, Luro F, Perrier X, Jacquemoud-Collet JP, Navarro L, Ollitrault P (2014) Next generation haplotyping to decipher nuclear genomic interspecific admixture in Citrus species: analysis of chromosome 2. BMC Genetics 15: 152

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Curk F, Ancillo G, Ollitrault F, Perrier X, Jacquemoud-Collet JP, Garcia-Lor A, Navarro L, Ollitrault P (2015) Nuclear Species-Diagnostic SNP Markers Mined from 454 Amplicon Sequencing Reveal Admixture Genomic Structure of Modern Citrus Varieties. Plos One 10: e0125628

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Curk F, Ollitrault F, Garcia-Lor A, Luro F, Navarro L, Ollitrault P (2016) Phylogenetic origin of limes and lemons revealed by cytoplasmic and nuclear markers. Annals of Botany 117: 565-583

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Espley RV, Hellens RP, Putterill J, Stevenson DE, Kutty-Amma S, Allan AC (2007) Red colouration in apple fruit is due to the activity of the MYB transcription factor, MdMYB10. Plant Journal 49: 414-427

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Fabroni S, Ballistreri G, Amenta M, Rapisarda P (2016) Anthocyanins in different Citrus species: an UHPLC-PDA-ESI/MSn-assisted qualitative and quantitative investigation. Journal of the Science of Food and Agriculture 96: 4797-4808

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Federici CT, Fang DQ, Scora RW, Roose ML (1998) Phylogenetic relationships within the genus Citrus (Rutaceae) and related genera as revealed by RFLP and RAPD analysis. Theoretical and Applied Genetics 96: 812-822

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Fincham JRS (1967) Mutable Genes in Light of Callans Hypothesis of Serially Repeated Gene Copies. Nature 215: 864-866

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Froelicher Y, Mouhaya W, Bassene JB, Costantino G, Kamiri M, Luro F, Morillon R, Ollitrault P (2011) New universal mitochondrial PCR markers reveal new information on maternal citrus phylogeny. Tree Genetics & Genomes 7: 49-61

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Garcia-Lor A, Ancillo G, Navarro L, Ollitrault P (2013) Citrus (Rutaceae) SNP Markers Based on Competitive Allele-Specific PCR; Transferability across the Aurantioideae Subfamily. Applications in Plant Sciences 1: 1200406

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Garcia-Lor A, Curk F, Snoussi-Trifa H, Morillon R, Ancillo G, Luro F, Navarro L, Ollitrault P (2013) A nuclear phylogenetic analysis: SNPs, indels and SSRs deliver new insights into the relationships in the 'true citrus fruit trees' group (Citrinae, Rutaceae) and the origin of cultivated species. Annals of Botany 111: 1-19

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Garcia-Lor A, Luro F, Ollitrault P, Navarro L (2015) Genetic diversity and population structure analysis of mandarin germplasm by nuclear, chloroplastic and mitochondrial markers. Tree Genetics & Genomes 11: 123

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Hall T (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 41: 95-98

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Heilmeyer M (2001) The Language of Flowers: Symbols and Myths. Prestel USA, New York

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Hillebrand S, Schwarz M, Winterhalter P (2004) Characterization of anthocyanins and pyranoanthocyanins from blood orange [Citrus sinensis (L.) Osbeck] juice. Journal of Agricultural and Food Chemistry 52: 7331-7338

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Hoballah ME, Gubitz T, Stuurman J, Broger L, Barone M, Mandel T, Dell'Olivo A, Arnold M, Kuhlemäier C (2007) Single gene-mediated shift in pollinator attraction in Petunia. Plant Cell 19: 779-790

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Hodgson RW (1967) History, world distribution, botany and varieties. In W Reuther, H Webber, L Batchelor, eds, The Citrus Industry. University of California Press, Berkeley, CA, pp 431-591

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Kimura M (1980) A Simple Method for Estimating Evolutionary Rates of Base Substitutions through Comparative Studies of Nucleotide-Sequences. Journal of Molecular Evolution 16: 111-120

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Kitajima A, Yamasaki A, Habu T, Preedasuttijit B, Hasegawa K (2007) Chromosome identification and karyotyping of satsuma mandarin by genomic in situ hybridization. Journal of the American Society for Horticultural Science 132: 836-841

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Kobayashi S, Goto-Yamamoto N, Hirochika H (2004) Retrotransposon-induced mutations in grape skin color. Science 304: 982

Pubmed: [Author and Title](#)

Downloaded from on October 5, 2020 - Published by www.plantphysiol.org
Copyright © 2017 American Society of Plant Biologists. All rights reserved.

CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Koes R, Verweij W, Quattrocchio F (2005) Flavonoids: a colorful model for the regulation and evolution of biochemical pathways. Trends in Plant Science 10: 236-242

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Krueger R, Navarro L (2007) Citrus germplasm resources. In IA Khan, ed, Citrus genetics, breeding, and biotechnology. CAB International, Wallingford, UK, pp 45-140

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Lai YT, Chen IZ (2008) Effect of Temperature on Calamondin (Citrus microcarpa) Flowering and Flower Bud Formation. Proceedings of the International Symposium on Citrus and Other Tropical and Subtropical Fruit Crops: 111-115

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Liu CY, Wang X, Xu YT, Deng XX, Xu Q (2014) Genome-wide analysis of the R2R3-MYB transcription factor gene family in sweet orange (Citrus sinensis). Molecular Biology Reports 41: 6769-6785

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Liu K, Warnow TJ, Holder MT, Nelesen SM, Yu JY, Stamatakis AP, Linder CR (2012) SATe-II: Very Fast and Accurate Simultaneous Estimation of Multiple Sequence Alignments and Phylogenetic Trees. Systematic Biology 61: 90-106

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Malinowski E (1935) Studies on Unstable characters in Petunia. The extreme flower types of the Unstable race with mosaic color patterns. Genetics 20: 342-356

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Matus JT, Aquea F, Arce-Johnson P (2008) Analysis of the grape MYB R2R3 subfamily reveals expanded wine quality-related clades and conserved gene structure organization across Vitis and Arabidopsis genomes. BMC Plant Biology 8: 83

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

McClintock B (1984) The Significance of Responses of the Genome to Challenge. Science 226: 792-801

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Moore A, Garibaldi C (2003) Flower Power: The Meaning of Flowers in Art 1500-2000. Philip Wilson Publishers, London

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Moore GA (2001) Oranges and lemons: clues to the taxonomy of Citrus from molecular markers. Trends in Genetics 17: 536-540

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Moore T (1866) The Treasury of Botany: a popular dictionary of the vegetable kingdom, Vol 1. Longmans Green & Co., London

Morton JF (1987) Grapefruit. In Fruits of warm climates. University of Florida, Miami, pp 152-158

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Netzel M, Netzel G, Tian QG, Schwartz S, Konczak I (2006) Sources of antioxidant activity in Australian native fruits. Identification and quantification of anthocyanins. Journal of Agricultural and Food Chemistry 54: 9820-9826

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Nicolosi E, Deng ZN, Gentile A, La Malfa S, Continella G, Tribulato E (2000) Citrus phylogeny and genetic origin of important species as investigated by molecular markers. Theoretical and Applied Genetics 100: 1155-1166

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Penjor T, Anai T, Nagano Y, Matsumoto R, Yamamoto M (2010) Phylogenetic relationships of Citrus and its relatives based on rbcL gene sequences. Tree Genetics & Genomes 6: 931-939

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Penjor T, Yamamoto M, Uehara M, Ide M, Matsumoto N, Matsumoto R, Nagano Y (2013) Phylogenetic Relationships of Citrus and Its Relatives Based on matK Gene Sequences. Plos One 8: e62574

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Quattrocchio F, Wing JF, Leppen HTC, Mol JNM, Koes RE (1993) Regulatory Genes-Controlling Anthocyanin Pigmentation Are Functionally Conserved among Plant-Species and Have Distinct Sets of Target Genes. Plant Cell 5: 1497-1512

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Quattrocchio F, Wing J, van der Woude K, Souer E, de Vetten N, Mol J, Koes R (1999) Molecular analysis of the anthocyanin2 gene of petunia and its role in the evolution of flower color. Plant Cell 11: 1433-1444

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Ramadugu C, Pfeil BE, Keremane ML, Lee RF, Maureira-Butler IJ, Roose ML (2013) A Six Nuclear Gene Phylogeny of Citrus (Rutaceae) Taking into Account Hybridization and Lineage Sorting. Plos One 8: e68410

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Ramsay NA, Glover BJ (2005) MYB-bHLH-WD40 protein complex and the evolution of cellular diversity. Trends in Plant Science 10: 63-70

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Sainsbury F, Thuenemann EC, Lomonossoff GP (2009) pEAQ: versatile expression vectors for easy and quick transient expression of heterologous proteins in plants. Plant Biotechnology Journal 7: 682-693

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Sanchez R, Serra F, Tarraga J, Medina I, Carbonell J, Pulido L, de Maria A, Capella-Gutierrez S, Huerta-Cepas J, Gabaldon T, Dopazo J, Dopazo H (2011) Phylemon 2.0: a suite of web-tools for molecular evolution, phylogenetics, phylogenomics and hypotheses testing. Nucleic Acids Research 39: W470-W474

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

SanMiguel P, Gaut BS, Tikhonov A, Nakajima Y, Bennetzen JL (1998) The paleontology of intergene retrotransposons of maize. Nature Genetics 20: 43-45

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Schwinn K, Venail J, Shang YJ, Mackay S, Alm V, Butelli E, Oyama R, Bailey P, Davies K, Martin C (2006) A small family of MYB-regulatory genes controls floral pigmentation intensity and patterning in the genus Antirrhinum. Plant Cell 18: 831-851

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Scora RW (1975) Symposium on Biochemical Systematics, Genetics and Origin of Cultivated Plants .9. History and Origin of Citrus. Bulletin of the Torrey Botanical Club 102: 369-375

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Steyn WJ, Wand SJE, Holcroft DM, Jacobs G (2002) Anthocyanins in vegetative tissues: a proposed unified function in photoprotection. New Phytologist 155: 349-361

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Stracke R, Werber M, Weisshaar B (2001) The R2R3-MYB gene family in Arabidopsis thaliana. Current Opinion in Plant Biology 4: 447-456

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Swingle WT (1946) The botany of Citrus and its wild relatives of the orange subfamily (family Rutaceae, subfamily Aurantioideae). In W Reuther, H Webber, L Batchelor, eds, The Citrus Industry. University of California Press, Berkeley, CA, pp 129-474

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Swingle WT, Reece PC (1967) The botany of Citrus and its wild relatives. In W Reuther, H Webber, L Batchelor, eds, The Citrus Industry. University of California Press, Berkeley, CA, pp 190-430

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Tamura K (1992) The rate and pattern of nucleotide substitution in Drosophila mitochondrial DNA. Mol Biol Evol 9: 814-825

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Molecular Biology and Evolution 28: 2731-2739

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Tanaka T (1977) Fundamental discussion of citrus classification. Studia Citrologica 14: 1-6

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Tanaka Y, Sasaki N, Ohmiya A (2008) Biosynthesis of plant pigments: anthocyanins, betalains and carotenoids. Plant Journal 54: 733-749

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Webber HJ (1943) Plant characteristics and climatology. In W Reuther, H Webber, L Batchelor, eds, The Citrus Industry, Vol 1. University of California Press, Berkeley, CA, pp 41-69

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Winkel-Shirley B (2001) Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. Plant Physiology 126: 485-493

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Wu GA, Prochnik S, Jenkins J, Salse J, Hellsten U, Murat F, Perrier X, Ruiz M, Scalabrin S, Terol J et al. (2014) Sequencing of diverse mandarin, pummelo and orange genomes reveals complex history of admixture during citrus domestication. Nature Biotechnology 32: 656-662

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Xu Q, Chen LL, Ruan XA, Chen DJ, Zhu AD, Chen CL, Bertrand D, Jiao WB, Hao BH, Lyon MP et al. (2013) The draft genome of sweet orange (Citrus sinensis). Nature Genetics 45: 59-U92

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)